submitted to *Phys. Chem. Chem. Phys.*

(Themed Issue: Advances in Mass Spectrometry for Biological Science)

Guided Ion Beam and Theoretical Studies of Sequential Bond Energies of Water to Sodium Cysteine Cation

Sha Joshua Ye and P. B. Armentrout*

Department of Chemistry, University of Utah, 315 S. 1400 E. Rm 2020, Salt Lake City, Utah 84112

Abstract

Absolute bond dissociation energies of water to sodium cysteine (Cys) cations and cysteine to hydrated sodium cations are determined experimentally by collision-induced dissociation of $Na^+Cys(H_2O)_x$, where x = 1 – 4, complexes with xenon in a guided ion beam mass spectrometer. Experimental results show that the binding energies of water and cysteine to the complexes decrease monotonically with increasing number of water molecules. Quantum chemical calculations at three different levels show reasonable agreement with the experimental bond energies. The calculations indicate that the primary binding site for Na^+ changes from charge-solvated tridentate chelation at the amino nitrogen, carbonyl oxygen, and sulfur side-chain for x = 0 and 1 to the C terminus of zwitterionic cysteine for x = 4, whereas different levels of theory provide conflicting predictions for x = 2 and 3. The first solvent shell of Na^+Cys is found to be complete at four waters. This is fewer than needed for the aliphatic amino acid glycine, because the functionalized side-chain of Cys provides an internal solvation site, a binding motif that probably applies for most other functionalized amino acids.

UU IR Author Manuscript

1. Introduction

The structures of biological macromolecules in aqueous solution are stabilized by their interactions with water and ions, such as Na^+ and K^+.[1,2] In addition, alkali metals are known to play an important role in biological systems, including modulating enzyme activity[3] and are key components in maintaining the cell membrane potential.[4] The overall thermodynamic outcome of the multitude of these interactions potentially can be understood by careful examination of the intrinsic interactions on a pairwise basis. Gas-phase studies provide a means to quantitatively assess these interactions in systems small enough for meaningful comparisons to theory, thereby also providing benchmark data for larger, less experimentally tractable systems.

In an aqueous environment, all amino acids (AAs) are in their zwitterionic form, yet in the gas phase, their uncharged tautomers are intrinsically more stable. The stabilities of amino acids, small peptides, and their analogues when complexed with alkali metal cations in the gas phase have been studied extensively in the past decade both experimentally and theoretically.[5-31] These interactions are mainly electrostatic and therefore the strength decreases with increasing size of the alkali metal ion and increases with chelation by functionalized side chains.[31-33] Although the presence of a metal cation significantly decreases the energy gap between the zwitterionic and charge-solvated forms of metalated complexes, in many cases, the charge-solvated form of the metallated amino acid is still preferred. An interesting question that arises in this context is how many water molecules are needed such that the zwitterionic form becomes the ground state.

Much less work exists for interactions of metallated biological systems when solvated with water. William's group investigated the hydration of metalated valine (Val) using black-body infrared radiative dissociation (BIRD). They find that the interaction between water and metalated valine also decreases with the size of the alkali metal ion and that the second hydration energy is weaker than the first.[34-37] Equilibrium-based methods such as BIRD are generally limited in the range of systems that can be easily studied, whereas the threshold collision-induced dissociation (TCID) experiments utilized in our laboratory can explore a more extensive

range. Indeed, we have used TCID experiments to determine a monotonic decrease of water binding energies with increasing number of water molecules for the $Na^+Gly(H_2O)_x$ system, $x = 1 - 4$, Gly = glycine.[38,39] Because valine and glycine are aliphatic / hydrophobic amino acids, they behave similarly in the gas phase when ionized by a sodium ion, leading to comparable water binding energies. The sodium ion favors [N,CO] coordination to nonzwitterionic Val and Gly (charge-solvated structures), but changes to CO coordination when hydrated by two or more water molecules.[17,34,36,38,39] Calculations on the hydrated sodium cation Val and Gly systems show that charge-solvated structures are still favored with 3 and 4 water molecules, respectively, whereas zwitterionic forms are favored in aqueous solution.

Changes are expected for hydration of Na^+Pro, where Pro = proline, because the secondary amine group of proline makes it more basic than most other amino acids. Thus experiments and calculations show that sodium cation binds to the carboxylate acid group of zwitterionic proline.[10,18,21,24,39] Williams and co-workers investigated the hydration of sodiated proline analogues determining the first hydration energy of Na^+α–Me–Pro.[40] In our laboratory, TCID experiments were used to determine the first four hydration enthalpies and computations showed that the proline remains zwitterionic, with the first solvent shell of Na^+Pro being essentially complete at four waters.[38,39]

In the present studies, the absolute bond dissociation energies (BDEs) of water to sodium cysteine cation and cysteine (Cys) to hydrated sodium cations are measured using competitive TCID methods. Complementary structural information about these systems is obtained by theoretical studies. Unlike Val, Gly, or Pro, the functionalized side-chain of Cys provides an intramolecular site for solvating the alkali metal cation, which may facilitate zwitterion formation by stabilizing the protonated amine group, such that differences in the binding patterns are anticipated. In this respect, we believe that the present results should be models for most amino acids having functionalized side chains. This is the first study to characterize the hydration thermodynamics of a metallated amino acid containing a functionalized side chain.

UU IR Author Manuscript

2. Experimental and Computational Section

2.1 General Experimental Procedures

The instrument used to measure the cross sections for TCID of the hydrated sodium cysteine cation complexes is a guided ion beam tandem mass spectrometer (GIBMS), described previously in detail.[41-43] Sodium ions are generated in the ion source using a continuous direct current glow discharge where the cathode is a tantalum boat filled with sodium metal and the surrounding source chamber is grounded. Typical operating conditions of the discharge are 1.6 – 2.2 kV and 15 – 25 mA. The sodium cations (Na^+) produced are carried by a flow of inert buffer gas (~ 10% Ar in He) through a 1 m long flow tube at a rate of 4000 – 9000 standard cm^3/min (sccm). At 10 cm downstream from the discharge, the neutral cysteine ligand is introduced using a temperature controllable heated probe (120 – 160 °C). Water vapor at room temperature is then introduced about 50 cm from the discharge through a pressure control valve. The complex ions of interest are formed via three-body associative reactions of Na^+ with the cysteine and water ligands in the flow of He/Ar gas. The complex ions are thermalized to 300 K (the temperature of the flow tube) both vibrationally and rotationally by undergoing ~10^5 collisions with the buffer gases in the 1 meter long flow tube at a pressure of 0.3 – 0.9 Torr.[44-47] Therefore, the rovibrational internal energies of all complex ions when exiting the flow tube can be described by a Maxwell-Boltzmann distribution at 300 K. After exiting the flow tube, the ionic complexes are focused into a 66° magnetic momentum analyzer that selects a particular $Na^+Cys(H_2O)_x$ complex, where x = 1 – 4. The ions are decelerated to a well-defined and variable kinetic energy and injected into a radio frequency double octopole ion beam guide,[41,43,48,49] which passes through a collision cell (8.3 cm effective length) containing the neutral reactant (here, Xe gas). The octopole serves to radially trap all unreacted complex ions and product ions formed by reactions with the neutral gas. Ions are extracted from the end of the octopole, mass analyzed by a quadrupole mass filter, and efficiently detected with a 27 kV conversion dynode-secondary electron scintillation detector[50] interfaced with fast pulse counting electronics.

UU IR Author Manuscript

Ion intensities, measured as a function of collision energy, are converted to absolute cross sections as described previously.[41] The absolute uncertainties in cross section magnitudes are estimated to be ± 20% and the relative uncertainties are approximately ± 5%. Laboratory (lab) energies are converted to center-of-mass (CM) energies using the equation $E_{CM} = E_{lab} \times M/(M + m)$, where M and m are the reactant neutral and ion masses, respectively. All energies cited below are in the CM frame unless otherwise noted. The absolute energy scale and the corresponding full width at half-maximum (fwhm) of the ion beam kinetic energy distribution are determined by using the octopole as a retarding energy analyzer.[41] The energy spread is nearly Gaussian and has a typical fwhm of 0.2 – 0.4 eV (lab) in the present experiments.

Because of the effects of multiple collisions,[51] we observe a slight dependence on Xe pressure for the cross section of the first dissociation product and an obvious dependence for the secondary and higher products. In order to obtain data free from such effects (i.e., at rigorously single-collision conditions), we collected data at about 0.04, 0.08, and 0.15 mTorr. The cross sections are then extrapolated to zero reactant pressure prior to threshold analysis, as described previously.[51]

2.2. Dissociation Threshold Analysis

To determine threshold energies for endothermic reactions, cross sections are modeled using eq 1,[52,53]

$$\sigma_j(E) = (n\sigma_{0,j}/E)\Sigma g_i \int_{E_{0,j}-E_i}^{E} [k_j(E^*)/k_{tot}(E^*)]\{1 - e^{-k_{tot}(E^*)\tau}\}(E-\varepsilon)^{n-1} d(\varepsilon) \quad (1)$$

where $\sigma_{0,j}$ is an adjustable scaling parameter for channel j, n describes the energy deposition efficiency during collision,[43] E is the relative kinetic energy, $E_{0,j}$ represents the CID threshold energy at 0 K for channel j, τ is the experimental time for dissociation (~ 5×10^{-4} s in the extended dual octopole configuration as measured by time-of-flight studies[43]), ε is the energy transferred from translation into internal energy during the collision, and E^* is the internal energy of the energized molecule (EM) after the collision, i.e., $E^* = \varepsilon + E_i$. The summation is over the rovibrational states i of the reactant ion having energies, E_i, and relative populations g_i,

where $\Sigma g_i = 1$, where a Maxwell-Boltzmann distribution at 300 K is assumed. The term $k_j(E^*)$ is the unimolecular rate constant for dissociation to channel j. This rate constant and $k_{tot}(E^*)$ are defined by Rice-Ramsperger-Kassel-Marcus (RRKM) theory as in eq 2,[54]

$$k_{tot}(E^*) = \sum_j k_j(E^*) = \sum_j d_j N_j^{\dagger}(E^* - E_{0,j}) / h\rho(E^*) \quad (2)$$

where d_j is the reaction degeneracy (defined using the ratios of rotational symmetries of reactants and products), $N_j^{\dagger}(E^* - E_{0,j})$ is the sum of rovibrational states of the transition state (TS) for channel j at an energy $E^* - E_{0,j}$, and $\rho(E^*)$ is the density of states of the EM at the available energy, E^*. Vibrational frequencies and rotational constants needed to determine the internal energy distributions of the ionic reactants are taken from the quantum chemical calculations detailed in the next section. The Beyer-Swinehart-Stein-Rabinovitch algorithm[55-57] is used to evaluate the number and density of the rovibrational states.

The model of eq 1 explicitly utilizes RRKM theory to estimate the kinetic shift associated with the possibility that energized molecules may not dissociate during the time scale of the experiment.[52] To evaluate the rate constants in eq 2, sets of rovibrational frequencies for the EM and all TSs are required. Because the metal ligand interactions in $Na^+Cys(H_2O)_x$ are mainly electrostatic (ion-dipole, ion-quadrupole, and ion-induced dipole interactions), the most appropriate model for the TS is a loose association of the ion and neutral ligand fragments. Therefore, the TSs are treated as product-like, such that the TS frequencies are those of the dissociation products. The molecular parameters needed for the RRKM calculation are taken from the quantum chemical calculations detailed in the next section. The transitional frequencies are treated as rotors, a treatment that corresponds to a phase space limit (PSL), as described in detail elsewhere.[52,53] For $Na^+Cys(H_2O)_x$ complexes, the five transitional mode rotors have rotational constants equal to those of the $Na^+Cys(H_2O)_{x-1}$ and H_2O products or $Na^+(H_2O)_x$ and cysteine products. The two-dimensional (2-D) external rotations are treated adiabatically but with centrifugal effects included.[58] In the present work, the adiabatic 2-D rotational energy is treated using a statistical distribution with an explicit summation over all possible values of the rotational quantum number.

UU IR Author Manuscript

The model cross section of eq 1 is convoluted with the kinetic energy distribution of the reactants[41] and compared to the data. A nonlinear least-squares analysis is used to provide optimized values for $\sigma_{0,j}$, $E_{0,j}$, and n. The uncertainty associated with $E_{0,j}$ is estimated from the range of threshold values determined from different data sets with variations in vibrational frequencies (± 10% and a factor of two for the M^+–L modes) and in the parameter n, variations in τ by a factor of 2, and the uncertainty in the absolute energy scale, 0.05 eV (lab). Final optimized BDEs at 0 K are obtained by making two assumptions. First, there is no activation barrier in excess of the reaction endothermicity for the loss of ligands, which is generally true for ion-molecule reactions, especially those such as the heterolytic bond cleavages considered here.[59] Second, the measured threshold $E_{0,j}$ for dissociation is from ground state reactant to ground state ion products and neutral H_2O or cysteine ligands. Given the relatively long experimental time frame (~5×10^{-4} s), energized complexes should be able to rearrange to low-energy product conformations upon dissociation.

2.3. Computational Details

The systems we examine here have many low-lying conformations. A simulated annealing procedure using the AMBER program and the AMBER forcefield based on molecular mechanics[60] was used to search for possible stable structures in each system's conformational space. All possible structures identified in this way were further optimized using NWChem.[61] at the HF/3-21G level.[62,63] Unique structures for each system that are within about 30 kJ/mol of the lowest energy structure (30 – 50 for each complex) were further optimized using Gaussian 03W[64] at the B3LYP/6-31G(d) level[65,66] with the "loose" keyword to facilitate more rapid convergence (maximum step size of 0.01 au and an rms force of 0.0017 au). The 15 to 30 lowest energy structures obtained from this procedure were then chosen for higher-level geometry optimizations and frequency calculations using density functional theory (DFT) at the B3LYP/6-311+G(d,p) level.[67,68] When used in zero-point vibrational energy (ZPE) corrections, internal energy determinations, or for RRKM calculations, these vibrational frequencies are scaled by 0.98.[69] We have shown in previous work on the related $Na^+Gly(H_2O)_x$ systems that MP2(full)

and B3LYP calculations using a 6-31G(d,p) basis set gave almost identical structural and energy information.[38] Therefore, MP2(full)/6-311+G(d,p) geometry optimizations were not performed for the present systems. Single point energy (SPE) calculations were carried out for these optimized structures at the B3LYP, B3P86, and MP2(full) levels using the 6-311+G(2d,2p) basis set.[67] Basis set superposition errors (BSSE) for all theoretical bond energies were estimated using the full counterpoise (cp) method in Gaussian 03W.[64,70] Previous work[17,38,71,72] indicates that cp corrections on alkali metal systems are generally small for DFT calculations and we find this to be true here as well. Both B3LYP and B3P86 SPE calculations have cp corrections less than 5 kJ/mol whereas they are 7 – 20 kJ/mol for the MP2(full) SPE calculations.

3. Results

3.1.Cross Sections for Collision-Induced Dissociation

Experimental cross sections were obtained for the interaction of Xe with $Na^+Cys(H_2O)_x$, $x = 1 - 4$. Fig. 1 shows representative data for CID of these complexes. Over the energy ranges examined, the dominant dissociation process for all $Na^+Cys(H_2O)_x$ complexes is the loss of one water molecule in reaction 3.

$$Na^+Cys(H_2O)_x + Xe \rightarrow Na^+Cys(H_2O)_{x-1} + H_2O + Xe \quad (3)$$

The magnitudes of the cross sections for losing one water molecule from $Na^+Cys(H_2O)_x$ increase from $x = 1$ to $x = 3$, roughly in proportion to the number of waters. For $x = 4$, the cross section is relatively small, which may indicate a different water binding pattern for the fourth water compared to the first three. The primary cross sections rise rapidly at low energy, level off, then generally decline at higher energies because of further dissociation of the primary product, as indicated by subsequent losses of water molecules to form $Na^+Cys(H_2O)_y$, $y = 1 - 3$. The absolute intensities of the reactant ion beams were found to decrease with increasing solvation such that our sensitivity for CID of $Na^+Cys(H_2O)_4$ is relatively poor, leading to the noisier cross sections in Fig. 1d.

In addition to reaction 3, loss of the cysteine ligand is observed in the competitive reaction 4 at higher energies in all cases.

$$Na^+Cys(H_2O)_x + Xe \rightarrow Na^+(H_2O)_x + Cys + Xe \quad (4)$$

The $Na^+(H_2O)_x$ cross sections are much smaller than those of the $Na^+Cys(H_2O)_{x-1}$ products. They rise more slowly than those for losing water and have apparent thresholds that are more than 0.8 eV higher. This indicates that the interaction of Na^+ with cysteine is stronger than that with H_2O, consistent with previous measurements, $D_0(Na^+–Cys)$ = 177 ± 5 kJ/mol and $D_0(Na^+–H_2O)$ = 98 ± 8 kJ/mol.[30,73] Loss of water molecules from the $Na^+(H_2O)_x$ products eventually forming Na^+ were also found for x = 1 – 3. Such products were not collected for x = 4 because of the lower sensitivity. In no system was a ligand exchange process forming a Xe containing ion observed.

3.2. Theoretical Results

As described above, multiple low energy structures of $Na^+Cys(H_2O)_x$ were optimized at the B3LYP/6-311+G(d,p) level of theory. The optimized structures of the lowest energy conformers are displayed below with additional structures illustrated in Fig. S1 of the supporting information. The single point energy values including zero point energy (ZPE) corrections calculated at three different levels of theory relative to the lowest energy isomer are given in Table 1 (with a more complete table included in the supporting information, Table S1).

To identify the structures of the complex, we start with the nomenclature established previously for Na^+Gly,[17,22,30] where the notation in brackets describes the metal binding sites for each conformer. The xW before the brackets indicates the number of water molecules attached to Na^+Cys, and the notation in brackets after xW indicates the water binding site unless the water molecule simply binds to the sodium ion through the oxygen atom of water. The specification of the metal binding sites is followed by a description of the cysteine orientation given by a series of four dihedral angles. These dihedral angles start with the carboxylic acid hydrogen atom (or analogous proton attached to the NH_2 group for zwitterionic structures) to define the ∠HOCC dihedral angle and proceed along the molecule to the terminal hydrogen of the amino acid side chain, i.e., ∠OCCC, ∠CCCS, and ∠CCSH (designated as c = cis indicating angles < 50°, g =

UU IR Author Manuscript

gauche for angles between 50° and 135°, and t = trans for angles > 135°). Among these, there exist pairs of structures that involve different orientations of the sulfur hydrogen relative to the cysteine backbone (g_+ or g_- in the CCSH dihedral angle), which show little energy difference (~ 2 – 5 kJ/mol). Generally, we discuss and show only the lower energy conformer of each pair. Several important geometric parameters for the $Na^+Cys(H_2O)_x$ complexes are provided in the supporting information, Table S2. For comparison purposes, we optimized the low energy structures of Na^+Cys and determined relative SPEs at the same levels of theory used here, and include these results in Tables 1 and S2 as well. These results are virtually identical to those reported previously.[30]

3.3. Na^+Cys and $Na^+Cys(H_2O)$

As reported elsewhere for Na^+Cys,[10,19,30] Na^+ favors [N,CO,S] tridentate coordination to nonzwitterionic cysteine in the gas phase. The current calculations yield the same results, with [N,CO,S]-tggg$_+$ being the ground state (GS) structure, with a tggg$_-$ variant (in which the HS bond points in the opposite direction away from the COOH group) only 3 – 5 kJ/mol higher, Table 1. A bidentate [N,CO]-tcgg$_+$ conformer is 14 – 16 kJ/mol higher than the GS. The lowest zwitterionic structure, [CO_2^-]-ctgg$_+$, is predicted to be 7 – 12 kJ/mol higher in energy than the GS (with a ctgg$_-$ variant lying 4 – 5 kJ/mol higher). Its charge-solvated analogue, [COOH]-ctgg$_+$, lies another 16 – 17 kJ/mol higher (with its ctgg$_-$ variant being essentially isoenergetic). The transition state (TS) connecting these two structures actually lies below the [COOH] structure once zero point energies are included (Table 1), indicating that [COOH] should collapse to its zwitterionic analogue. Higher energy structures (>21 kJ/mol above the GS) were also located having [N,OH,S] and [CO,S] binding configurations, but these are never found to be among the lowest energy conformers of the hydrated complexes.

For $Na^+Cys(H_2O)$, all levels of theory (Table 1) predict the same GS, a charge-solvated isomer in which the water molecule attaches directly to Na^+ in the GS conformer of Na^+Cys to form 1W[N,CO,S]-tggg$_+$, such that the metal ion is four coordinate, Fig. 2. As for Na^+Cys, the structure with sulfur hydrogen pointing away from the carboxylic acid group (tggg$_-$) is only 3 –

5 kJ/mol higher. The two charge-solvated bidentate structures, 1W[N,CO]-tcgg and 1W[N,CO]-tgtg are 10 – 17 kJ/mol above the tridentate structures, 1 – 5 kJ/mol more stable than the difference between the bidentate and tridentate structures for Na^+Cys. The water can also bind to the hydroxyl group rather than Na^+, forming 1W[HO]-[N,CO,S]-tggg$_+$, but this structure lies 13 – 19 kJ/mol higher in energy.

The lowest zwitterionic structure, 1W[CO_2^-]-ctgg$_+$, lies 5 – 12 kJ/mol higher than the GS, with a variant, 1W[CO_2^-]-ctgg$_-$, lying 5 – 6 kJ/mol higher still, similar to the relative energies observed for Na^+Cys. An alternative zwitterionic structure, 1W[bO^-]-[CO_2^-]-ctgg$_+$, positions the water molecule such that it bridges between the sodium cation and the oxygen of the carboxylate group that is bound to the NH_3^+ group, Fig. 2. The water can also bridge to the other oxygen, forming 1W[bCO]-[CO_2^-]-ctgg$_+$. These two conformers lie 4 – 7 and 10 – 13 kJ/mol, respectively, above 1W[CO_2^-]-ctgg$_+$. We also located a zwitterionic structure in which the water molecule binds to the NH_3^+ group, 1W[HN]-[CO_2^-]-ctgg$_+$, but this lies 17 – 20 kJ/mol above the zwitterionic conformer where the water binds to the sodium cation.

The charge solvated analogue of the lowest zwitterionic structure is 1W[COOH]-ctgg$_+$, which lies 9 – 10 kJ/mol higher in energy. Likewise the 1W[bOH]-[CO]-ctgg$_+$ structure lies 7 – 8 kJ/mol above its zwitterionic analogue, 1W[bO^-]-[CO_2^-]-ctgg$_+$, with its ctgg$_-$ variant within 0.5 kJ/mol. We also located the TS for proton motion between the nitrogen in the 1W[CO_2^-]-ctgg$_+$ structure and the oxygen in the 1W[COOH]-ctgg$_+$ structure. This TS lies 3 kJ/mol below the latter structure after zero point energy corrections are included at the B3P86 level and about 1 – 2 kJ/mol above at the B3LYP and MP2 levels (Table 1). Thus, the [COOH] structure is stabilized by 2 – 3 kJ/mol by the addition of the water ligand.

In most of these structures, the addition of one water molecule to Na^+Cys has little effect on the original binding pattern of Na^+Cys (evident by nearly the same ∠Na^+OC bond angles and ∠Na^+OCC and ∠NCCO dihedral angles, Table S2). However, the binding distances between the metal ion and oxygen, nitrogen, and sulfur atoms are elongated slightly (by 0.01 – 0.04 Å) because of the electron delocalization from water to the sodium cation, which reduces the

UU IR Author Manuscript

electrostatic interaction with cysteine. Exceptions to this behavior include structures involving a bridging water, which allows the sodium ion to get closer to the carboxylate or carbonyl oxygen (by up to 0.07 Å) and permits more linear ∠NaOC bond angles (116° – 133° versus 89° – 98° in Na^+Cys). However, because the sodium cation no longer interacts directly with both oxygens as it does in 1W[CO_2^-]-ctgg$_+$, these bridging structures are higher in energy, Table 1 and Fig. 2. Structures in which the water binds to either the HO or HN site also exhibit small decreases in the sodium cation binding distances (0.01 – 0.02 Å), presumably because the water donates electron density to the ligand.

3.4. $Na^+Cys(H_2O)_2$

2W[N,CO,S]-tggg$_+$ is predicted to be the ground conformer by the MP2(full) method (with the tggg$_-$ variant still 4 – 6 kJ/mol higher), whereas the B3P86 approach predicts a zwitterionic 2W[bO^-]-[CO_2^-]-ctgg$_+$ GS, Fig. 3 (with ctgg$_-$ and cgtg variants 4 – 6 kJ/mol higher). B3LYP suggests these structures are isoenergetic (with the latter only 0.004 kJ/mol higher in energy). These structures can be regarded as further hydration of the sodium ion in 1W[N,CO,S]-tggg$_+$ and 1W[bO^-]-[CO_2^-]-ctgg$_+$, respectively. Note that the second water stabilizes this bridging zwitterionic conformer by 10 – 16 kJ/mol compared to the charge-solvated tridentate structure, which we believe is largely a consequence of the five-coordinate sodium cation in 2W[N,CO,S] being sterically crowded. This conclusion is substantiated by noting that another low-lying structure is 2W[HO]-[N,CO,S]-tggg$_+$ (Fig. 3). This species lies about 3 kJ/mol below (B3P86) to 6 kJ/mol above (MP2) the 2W[N,CO,S]-tggg$_+$ conformer, a stabilization of 13 – 18 kJ/mol compared to the analogous single water complexes. Thus, addition of the second water to the HO site (such that the complex retains four-coordinate Na^+) is thermodynamically competitive with addition at the Na^+ site yielding a sterically-crowded five-coordinate Na^+. The bidentate 2W[N,CO]-tcgg structure is stabilized by 3 – 7 kJ/mol by the addition of the second water, presumably because the solvation shell of Na^+ is better satisfied by four-coordination.

UU IR Author Manuscript

Another low-lying structure is 2W[bOH]-[CO]-ctgg$_+$ (with the ctgg$_-$ variant within 1 kJ/mol and a cgtg variant about 3 – 5 kJ/mol higher), which all three levels of theory indicate lies only 2 – 3 kJ/mol above its zwitterionic analogue, 2W[bO$^-$]-[CO$_2^-$]-ctgg$_+$. A zwitterionic variant where one water bridges to the carbonyl oxygen rather than the oxygen that is hydrogen-bound to NH_3^+ is 2W[bCO]-[CO$_2^-$]-ctgg$_+$, which theory locates 6 – 7 kJ/mol higher than the lowest energy zwitterion, the same difference observed for the 1W complexes. Binding one water directly to the NH_3^+ site results in a structure, 2W[HN]-[CO$_2^-$]-ctgg$_+$ (Fig. S1), lying 7 – 9 kJ/mol higher than 2W[bO$^-$]-[CO$_2^-$]-ctgg$_+$. (Here, the ctgg$_-$ variant lies another 5 – 7 kJ/mol higher.) If instead the second water is bridging, 2W[HN,bO$^-$]-[CO$_2^-$]-ctgg$_+$ (Fig. S1), the energy increases another 4 – 7 kJ/mol (with the ctgg$_-$ variant another 5 kJ/mol higher). The water can also bridge from Na^+ to the thiol side chain, forming 2W[bOH,bS]-[CO]-ctgt (Fig. S1). This species lies 12 – 15 kJ/mol above the lowest energy zwitterion, with its zwitterionic analogue, 2W[bO$^-$,bS]-[CO$_2^-$]-cggt, another 16 – 18 kJ/mol higher.

The transition state connecting the GS zwitterion, 2W[bO$^-$]-[CO$_2^-$]-ctgg$_+$, with its charge-solvated analogue, 2W[bOH]-[CO]-ctgg$_+$, is 3 – 4 kJ/mol above the latter structure when ZPE corrections are included at the B3LYP and MP2 levels (~7 kJ/mol for the ctgg$_-$ variant), but 1.4 kJ/mol below at the B3P86 level (0.6 kJ/mol above for ctgg$_-$). All three levels agree that the energy variation among these ctgg structures is small (<6 kJ/mol for ctgg$_+$ and <10 kJ/mol for ctgg$_-$).

Addition of the second water molecule to the Na^+ site increases the sodium-ligand bonds rather uniformly by 0.03 – 0.05 Å, when retaining the binding pattern of Na^+ to Cys. As for addition of a single water, if the second water binds in a bridging position or to the HO or HN positions, then the sodium-ligand bonds decrease by 0.01 – 0.07 Å. Other geometric variations follow the same patterns noted above for the 1W complexes.

3.5. $Na^+Cys(H_2O)_3$

For the three water complex, the lowest energy tridentate structure is 3W[HO,bCO]-[N,CO,S]-tggg$_+$ (Fig. 4), which has an elongated Na^+-OC bond length because of the bridging

UU IR Author Manuscript

water. Although low-lying, the tridentate structure is no longer the lowest energy conformer, lying 1 – 8 kJ/mol above the GS, which is predicted differently by each level of theory. MP2 and B3LYP calculations find a 3W[bOH,bS]-[CO]-ctgt GS in which the sodium cation is tetracoordinate, binding to the Cys carbonyl and three waters with one water bridging to the hydroxyl group and another water bridging to the thiol group. This species can be thought of as addition of a water bridging to the HS of the 2W[bOH]-[CO]-ctgg$_-$ conformation. B3P86 calculations also find this structure to be low-lying, 3 kJ/mol above the GS, Table 1. Notably, this charge-solvated structure does not have a direct zwitterionic analogue, as moving the proton from OH···N to O···HN forms 3W[bO$^-$,bW]-[CO_2^-]-ctgg$_-$, 10 – 14 kJ/mol higher in energy, in which the hydrogen bond to the thiol group has been replaced by a hydrogen bond between two of the water molecules. The TS connecting these two structures lies 1.5 kJ/mol below (MP2) to 3 kJ/mol above (B3LYP) the zwitterion.

Addition of a water molecule to the Na^+ or HN sites of 2W[bO$^-$,bW]-[CO_2^-]-ctgg$_+$, leads to two low energy zwitterions. 3W[HN,bO$^-$]-[CO_2^-]-ctgg$_+$ (Fig. 4), which B3P86 indicates is the GS, is only 1 – 5 kJ/mol lower than 3W[bO$^-$,bW]-[CO_2^-]-ctgg$_+$, illustrating a near isoenergetic competition between the Na^+ and HN hydration sites. For the former zwitterion, the proton motion leads to the charge-solvated form, 3W[HN,bOH]-[CO]-ctgg$_+$, which is 20 – 22 kJ/mol above the zwitterion. These species have a connecting TS lying 2 – 7 kJ/mol below the charge-solvated structure once zero point energies are included, indicating the charge solvated form will collapse to the zwitterion. For the 3W[bO$^-$,bW]-[CO_2^-]-ctgg$_+$ zwitterion, the charge-solvated analogue, 3W[bOH]-[CO]-ctgg$_+$ is the B3LYP GS and lies 2 – 5 kJ/mol lower in energy. Their connecting TS lies 0 – 4 kJ/mol above the zwitterionic form.

With three water molecules, all low-lying structures (<30 kJ/mol) have at least one bridging water ligand and binding to the HO or HN positions is comparable in energy to binding to the sodium cation. Trends in the bond distances noted above for the 1W and 2W complexes continue here. Several other structures were also located as listed in Tables 1 and S1 and lie 4 – 32 kJ/mol above the GS structures. The relative stabilities of bidentate [N,CO] structures, e.g.,

3W[bS,bCO,bW]-[N,CO]-tggg$_+$, Fig. 4, are not well characterized by theory, with the third water stabilizing such structures by 1 – 5 kJ/mol at the B3LYP and MP2 levels and having no effect at the B3P86 level.

3.6. $Na^+Cys(H_2O)_4$

By the time the fourth water molecule is added to the Na^+Cys complex, the system is completely saturated. All three levels of theory agree that the lowest energy structure is the zwitterionic 4W[HN,bO$^-$,bW]-[CO_2^-]-ctgg$_+$, Fig. 5 and Table 1. (Variants, cgtg$_-$ and ctgg$_-$, lie 4 – 6 and 6 – 8 kJ/mol higher in energy, respectively.) Here the NH_3^+ site hydrogen bonds to the carboxylate, the thiol group, and a water such that it is fully solvated. The Na^+ site is tetracoordinate, binding to a carboxylate oxygen and three water molecules, one of which bridges to the carboxylate and one to another water ligand. The charge-solvated analogue of this complex, 4W[HN,bOH]-[CO]-ctgg$_+$, lies much higher in energy, by 16 – 19 kJ/mol, with the TS connecting them 0 – 5 kJ/mol lower after zero point energy corrections, Table 1 and S1. Thus this charge-solvated structure collapses to the zwitterion at thermal energies. The lowest-lying charge-solvated conformation, only 0.2 – 3.4 kJ/mol above the GS, is 4W[bOH,AA]-[CO]-ctgg$_+$, in which the fourth water molecule (which we designate as an acceptor-acceptor, AA) binds in a second solvent shell to two waters that are bound directly to Na^+. (A cgtg$_-$ variant lies 4 – 5 kJ/mol higher.) The zwitterionic analogue of this complex, 4W[bO$^-$,AA]-[CO_2^-]-ctgg$_+$, lies 4 – 6 kJ/mol higher in energy. Other structures listed in Tables 1 and S1 generally follow the same patterns noted above. In the case of the 4W[HN-bW,bO$^-$,AD]-[CO_2^-]-cgtg$_-$ complex, lying 4 – 10 kJ/mol above the GS, one water binds to the HN position while bridging to a water bound to Na^+ (Fig. 5). This latter water also hydrogen bonds to the water bridging Na^+ and the carboxylate, such that it has both an acceptor and donor interaction (AD) with other water ligands. Similarly, in two 4W[HN-bO$^-$,bO$^-$,bW]-[CO_2^-]-cgtg$_-$ structures, lying 6 – 10 kJ/mol above the GS, a water bound to the NH_3^+ group bridges to a carboxylate oxygen. There are two nearly isoenergetic variants in which the HN-bO$^-$ and bW waters bind either on the same side of the complex as the side-chain (r = rear) or on the opposite side (f = front, Fig. 5).

UU IR Author Manuscript

The lowest energy isomer in which Na^+ retains a bond to the amine is 4W[HO,bS,bCO,bW]-[N,CO]-tggg$_+$. Again the sodium cation is five-coordinate and there are extensive bridging interactions. Moving the water bound to the hydroxy group to the sodium, forming 4W[bS,bCO,bHN,bW]-[N,CO]-tggg$_+$ (Fig. S1), increases the energy by 3 (MP2) to 12 (DFT) kJ/mol because the sodium cation is now six-coordinate, and the cysteine distorts to accommodate such a crowded coordination, Table S2. The final structure shown in Fig. 5, 4W[AA-bCO]-[CO_2^-]-ctgg$_+$, shows another interesting binding motif in which the second solvent shell water (AA) bridges to the carboxylate, which again allows for a relatively short Na^+-OC bond length, Table S2. This structure lies 10 – 12 kJ/mol above the GS.

3.7. Relative free energies at 298 K

In considering what species are formed in our thermal ion source, the relative 298 K free energies are more relevant than 0 K enthalpies. Conversion of the calculated 0 K energies to 298 K free energies is accomplished using the rigid rotor/harmonic oscillator (RR/HO) approximation and the frequencies calculated at the B3LYP/6-311+G(d,p) level. These relative ΔG_{298} values are also reported in Table 1. It should be recognized that for floppy systems such as those examined here, the use of the RR/HO approximation for all modes may be inaccurate as many of the water ligands have torsional modes that would be better represented as hindered or free rotors. An accurate estimate of such motions is beyond the scope of the present study.

For the x = 0 and 1 complexes, the GS species determined at 0 K remain the ground states at 298 K, such that the tridentate [N,CO,S] structures should dominate the species formed. For the x = 2 complexes, 2W[N,CO,S] remains the 298 K GS at the B3LYP and MP2 levels, but the B3P86 GS changes from 2W[bO^-]-[CO_2^-] to the charge-solvated 2W[bOH]-[CO]-ctgg$_-$, which is also low-lying at the B3LYP level (but 9.8 kJ/mol above the GS at the MP2 level), Table 1. The B3P86 GS changes because these two conformers show a modest difference in $\Delta G_{298} - \Delta H_0$ (~5 kJ/mol), which primarily reflects the fact that zwitterionic structures have higher vibrational frequencies than analogous charge-solvated structures, presumably because of the more highly localized charges. Another factor that influences the free energies is a bridging water ligand has

UU IR Author Manuscript

much more restricted motions than non-bridged waters. For example, there are two 2W[HN]-[CO_2^-] structures that drop considerably in 298 K free energy because neither water is bridging.

For the x = 3 complexes, the 3W[bOH,bS]-[CO] structure (the 0 K B3LYP and MP2 GS) increases considerably in 298 K free energy because two of the water ligands are bridging. With only one bridging ligand, 3W[HO,bCO]-[N,CO,S] becomes the MP2 298 K free energy GS. At the two DFT levels of theory, the 298 K free energy GS becomes the 3W[bOH]-[CO] conformer, which is also low in free energy at the MP2 level, Table 1. For the x = 4 complexes, the $\Delta G_{298} - \Delta H_0$ difference for the 0 K GS, 4W[HN,bO^-,bW]-[CO_2^-], is relatively low such that it remains the 298 K free energy GS at all three levels of theory. The only complex with an appreciably smaller $\Delta G_{298} - \Delta H_0$ value is 4W[HN,bOH]-[CO], with only a single bridging water; however, this species has a sufficiently high ΔH_0 (16 – 19 kJ/mol) that it remains relatively high lying at 298 K (8 – 11 kJ/mol), Table S1.

3.8. Threshold Analysis and Results

We have shown previously[24,53,74] that the best measurements of the thresholds for competitive dissociation processes come from the simultaneous analysis of the cross sections. In general, the apparent threshold for the higher energy process is elevated from its thermodynamic value because of competition with the lower energy channel. In the present $Na^+Cys(H_2O)_x$ systems, the competitive model of eq 1 was used to analyze the competitive processes 3 and 4 for $Na^+Cys(H_2O)_x$ systems with x = 1 – 4. Fig. 1 shows that all experimental cross sections for reactions 3 and 4 are reproduced well by eq 1 over energy ranges exceeding 3 eV. The optimized parameters of eq 1 are reported in Table 2 along with results from previous work on CID of the Na^+Cys complex.[30]

In previous studies of competitive dissociations,[38,53,74] we found that independent scaling factors (different values of $\sigma_{0,j}$ for each channel) are sometimes needed in order to reproduce the experimental data when using the competitive analysis. The use of independent scaling factors compensates for neglected factors, such as reaction degeneracies, symmetry numbers of the reactant and product molecules, dipole moments of neutral products, and inaccurate estimations

UU IR Author Manuscript

of metal ligand frequencies, although all of these factors are included in the modeling to the best of our ability to estimate them. In addition, hindered rotors and their symmetries can also influence such scaling,[74] but were not explicitly included here. In the present study, independent scaling factors were needed only for $x = 1$ with a relative $\sigma_{0,j}$ value of ~27, whereas common scaling factors could be used for $x = 2 - 4$ (Table 2). In the latter cases, it would appear that all physically relevant factors are adequately represented in the approach used here.

The experimental results in Table 2 were obtained using molecular parameters for the 298 K free energy GS reactants dissociating to 0 K GS products, as calculated at the B3LYP/6-311+G(d,p) level in all cases. This is because the 298 K free energy GS species should represent the dominant complexes formed in the thermal ion source, whereas we have shown that thresholds obtained using eq 1 are associated with formation of products at 0 K.[75] For $x = 2 - 4$, we checked whether using different molecular structures for the $Na^+Cys(H_2O)_x$ complexes and their $Na^+Cys(H_2O)_{x-1}$ products changed any of the fitting parameters, especially because there are multiple structures predicted to be the GS structure by different theoretical methods for $x = 2$ and 3. In nearly all cases, values for $\sigma_{0,j}$, n, and $E_{0,j}$ were nearly unchanged (threshold energy differences < 0.02 eV), with the only exception being for $x = 4$. Here, dissociation of 4W[HN,bO$^-$,bW]-[CO_2^-] to 3W[bOH]-[CO] gave a lower threshold for Cys loss by 0.09 ± 0.01 eV compared to formation of 3W[bOH,bS]-[CO]. A factor that did change appreciably in some cases was the entropy of activation, $\Delta S^{\ddagger}$, which characterizes the looseness of the transition states involved (here the product channel). For all complexes, the $\Delta S^{\ddagger}$ values for losing water or cysteine increase with x, Table 2. No matter what molecular parameters we use, the entropy of activation for water loss is much smaller, by 40 – 70 J/(mol·K), than for cysteine loss, which reflects the constraints on the torsional motions of cysteine when complexed to the sodium cation.

3.9. Thermodynamic cycles

For the competitive modeling results, the accuracy of the values can be checked by comparing to values obtained using the thermodynamic cycle shown in Fig. 6. For instance, the energies required for loss of Cys from each hydrated complex can be calculated using eq 5,

$$E_0[Na^+Cys(H_2O)_x \rightarrow Cys + Na^+(H_2O)_x] = E_0[Na^+Cys(H_2O)_x \rightarrow Na^+Cys(H_2O)_{x-1} + H_2O] + E_0[Na^+Cys(H_2O)_{x-1} \rightarrow Cys + Na^+(H_2O)_{x-1}] - E_0[Na^+(H_2O)_x \rightarrow Na^+(H_2O)_{x-1} + H_2O] \quad (5)$$

where the first two E_0 values on the right of the equation are measured here (except for the $x = 0$ value)[30,73] and the last elsewhere.[30,73] Thus, for any complex, the energy required to lose cysteine equals the energy of removing one water, then cysteine from $Na^+Cys(H_2O)_{x-1}$, minus the energy gained by adding the water back to hydrated Na^+. The results from these calculations are listed in Table 3 (Cycle 1) and compared with the threshold values obtained by competitive modeling, where good agreement is found in all cases.

Two other related thermodynamic cycles can also be envisioned. Cycle 2 is anchored to $D_0(Na^+–Cys)$ and values for loss of Cys from the hydrated complexes are obtained by removing all the waters, then losing Cys from Na^+Cys, then adding the waters back to bare Na^+, eq 6.

$$E_0[Na^+Cys(H_2O)_x \rightarrow Na^+(H_2O)_x + Cys] = E_0[Na^+Cys(H_2O)_x \rightarrow Na^+Cys + xH_2O] + E_0[Na^+Cys \rightarrow Na^+ + Cys] - E_0[Na^+(H_2O)_x \rightarrow Na^+ + xH_2O] \quad (6)$$

Alternatively, we can use the value for the largest complex ($x = 4$) as the anchor point, cycle 3.

$$E_0[Na^+Cys(H_2O)_x \rightarrow Na^+(H_2O)_x + Cys] = E_0[Na^+Cys(H_2O)_4 \rightarrow Na^+(H_2O)_4 + Cys] + E_0[Na^+(H_2O)_4 \rightarrow Na^+(H_2O)_x + (4-x)H_2O] - E_0[Na^+Cys(H_2O)_4 \rightarrow Na^+Cys(H_2O)_x + (4-x)H_2O] \quad (7)$$

These values along with the weighted average of the $Na^+(H_2O)_x$–Cys bond energies obtained from the various thermodynamic cycles and competitive fits (where the uncertainties listed are two standard deviations of the mean) are reported in Table 3. These average values are our best determinations of this thermochemistry and used throughout the remainder of the paper along with bond energies for water loss taken from the competitive E_0 results in Table 2. Overall, these comparisons of these various results confirm that the BDEs recommended in the present study are self-consistent and compatible with previous work in our laboratory.[30,73]

3.10. Conversion of binding energies from 0 to 298 K

We also convert our experimental 0 K bond energies to 298 K bond enthalpies and free energies by using the rigid rotor/harmonic oscillator (RR/HO) approximation and the frequencies

UU IR Author Manuscript

calculated at the B3LYP/6-311+G(d,p) level. These ΔH_{298} and ΔG_{298} values along with the conversion factors and the 0 K enthalpies are reported in Table 4. Uncertainties listed include variations associated with scaling most vibrational frequencies by ±10% and the metal-ligand frequencies by factors of two. Because theory provides different ground state structures for x = 2 and 3, the uncertainties also include the range of values obtained by considering all possible ground state species as reactants and products. As above, the use of the RR/HO approximation for all modes may be inaccurate because of the very low frequencies associated with torsional modes of the water ligands.

4. Discussion

4.1. Trends in experimental bond dissociation energies

Our best results for the energies required to remove water and cysteine from $Na^+Cys(H_2O)_x$, x = 1 – 4, are shown in Fig. 7 (competitive analysis thresholds for losing water, Table 2, and the weighted average for losing cysteine from Table 3) along with previous experimental results for $D_0(Na^+–Cys)$.[30] The BDE for losing cysteine is larger than that for losing water from each $Na^+Cys(H_2O)_x$ complex, x = 1 – 4, consistent with the qualitative dissociation behavior exhibited in Fig. 1. Similar to the $Na^+(H_2O)_x$ system,[73] the BDEs of water to sodiated cysteine decrease with increasing number of water molecules because of increasing steric effects and decreasing effective charge on the sodium ion. The first and second water BDEs to sodiated cysteine (66 ± 9 and 54 ± 8 kJ/mol) are very similar to those of the third and fourth water to the sodium cation (70 ± 6 and 55 ± 6 kJ/mol) in the $Na^+(H_2O)_x$ system.[73] This correspondence implies that the solvation effect of cysteine on a sodium ion is comparable to that of two water molecules. Indeed, the BDE of cysteine to sodium ion is 177 ± 5 kJ/mol,[30] the same as the sum of the BDEs for the first and second water on Na^+ (177 ± 10 kJ/mol).[73] In addition, experimental BDEs of cysteine to $Na^+(H_2O)_x$ also decrease with increasing solvation from x = 0 to 4, Fig. 7.

UU IR Author Manuscript

The experimental binding energy of the fourth water molecule is measured to be 34 ± 7 kJ/mol, 11 kJ/mol larger than the energy associated with the hydrogen bonding network in pure water, 23 kJ/mol.[76] This result implies that the fourth water is still binding to a favorable position on cysteine or Na^+. Indeed, all levels of theory predict the same zwitterionic ground isomer of $Na^+Cys(H_2O)_4$, 4W[HN,bO$^-$,bW]-[CO_2^-]-ctgg$_+$, Fig. 5, where one of the water molecules hydrogen bonds to $-NH_3^+$ and Na^+ is tetracoordinate with its first solvent shell formed by the carbonyl oxygen and the other three water molecules. One imagines that any additional water molecules will hydrogen bond primarily to the water molecules already present, leading to BDEs that are somewhat weaker than that of the fourth water ligand bond, while continuing to favor the zwitterionic form of cysteine. The beginnings of such conformations can be seen in the 4W[bOH,AA]-[CO]-ctgg$_+$, 4W[HN-bW,bO$^-$,AD]-[CO_2^-]-cgtg$_-$, and 4W[AA-bCO]-[CO_2^-]-ctgg$_+$ structures, which lie within 12 kJ/mol of the GS. Furthermore, the hydration energies measured here decrease by 11 ± 1 kJ/mol for addition of each water ligand, such that the fifth water molecule is predicted to have a bond energy comparable to that for pure water-water interactions.

4.2. Comparison of Theoretical and Experimental Bond Dissociation Energies

Previously,[39] we performed calculations for $Na^+(H_2O)_x$, x = 1 – 4, using the same procedure outlined above for $Na^+Cys(H_2O)_x$. This yields BDEs (in kJ/mol) at the B3LYP, B3P86, and MP2 levels of theory of 94.3, 91.0, and 88.9 for x = 1; 82.8, 79.9, and 78.6 for x = 2; 65.0, 63.0, and 63.5 for x = 3; and 53.0, 50.3, and 53.3 for x = 4. Almost all the calculated values agree with experimental values within uncertainties,[73] with mean absolute deviations (MADs) of 2 ± 2, 4 ± 2, and 4 ± 2 kJ/mol, respectively.

The theoretical BDEs for the $Na^+Cys(H_2O)_x$ complexes calculated at three levels of theory are compared to the experimental values in Table 5. We find that the three theoretical methods (B3LYP, B3P86, and MP2) yield BDE values for the $Na^+Cys(H_2O)_x$ systems that can differ by up to 10 kJ/mol from one another, but overall show reasonable agreement with our experimental values, Fig. 7. MADs from experiment are 3 – 4 kJ/mol for losing water and 7 – 8 kJ/mol for losing cysteine, Table 5. The MP2(full) results give marginally the best comparison to

experiment, as also found recently for hydration energies of Zn^{2+}.[75] Similar to the $Na^+(H_2O)_x$ results, the theoretical values for loss of water generally agree with experiment within experimental uncertainty. The calculated BDEs for losing cysteine from the $Na^+Cys(H_2O)_x$ complexes also show agreement with the experimental values within the larger experimental uncertainties, although have larger MADs compared to the water loss channels. Comparison of the various cycles in Table 3 indicate that Cycle 2 (anchored to Na^+Cys), Cycle 3 (anchored to the $x = 4$ value), and the average give the best agreement with theory (with MADs of 6 kJ/mol), but only slightly better than the competitive fit or Cycle 1 values. Overall, the agreement between experimental and theoretical BDEs is quite reasonable with both following the same trends, Fig. 7.

4.3. Influence of the Side-chain on the Hydration of Amino Acids

The experimental BDEs of hydrated Na^+Gly[38] and Na^+Pro[39] are compared to those of Na^+Cys in Fig. 8. In all three systems, the BDEs for losing the amino acid (AA) from $Na^+AA(H_2O)_x$, $x = 0 - 4$, are higher than the BDEs for losing water at each x. BDEs for losing water from $Na^+Pro(H_2O)_x$ are 9 – 12 kJ/mol lower than the corresponding BDEs for $Na^+Gly(H_2O)_x$ system at all x, whereas those for $Na^+Cys(H_2O)_x$ start off as intermediate ($x = 1$ and 2) and become slightly larger for $x = 3$ and 4. For the smaller complexes, the trends reflect the relative binding energies of sodium cations to these three amino acids, Gly < Cys < Pro,[17,24,30] such that the more tightly bound amino acid binds additional water molecules less tightly. This can also be seen in the bond energies for loss of AA from the complexes of $x = 0 - 2$.

Theoretical calculations show that the ground state structures are nonzwitterionic for $Na^+Gly(H_2O)_x$, $x = 0 - 4$, zwitterionic for $Na^+Pro(H_2O)_x$, $x = 0 - 4$, and evolving for $Na^+Cys(H_2O)_x$, $x = 0 - 4$. Na^+ has bidentate binding to glycine with a [N,CO] configuration, bidentate to proline at the CO_2^- group, and tridentate to cysteine, [N,CO,S]. The first water binds directly to Na^+ for all three amino acids. Upon addition of a second water, the most favorable binding sites for Na^+ to Gly and Pro are both CO coordination with one water molecule bridging

to the hydroxyl oxygen,[38,39] whereas the [N,CO,S] tridentate structure is essentially isoenergetic with this bridging conformation for Cys. Thus, as waters are added to Na^+Gly, the binding site changes to chelation at the C terminus of glycine ([COOH] coordination) for $x = 2 - 4$.[38] In the Gly system, all water molecules for $x = 1 - 4$ directly solvate Na^+, whereas for the Pro and Cys systems, solvation of the protonated amino group become comparable in energy to direct Na^+ solvation for three and four water molecule systems. In the $Na^+Cys(H_2O)_x$ system, solvation of the hydroxyl group is also possible, whereas the Pro system remains zwitterionic throughout. For $Na^+Cys(H_2O)_4$, the structure where both $-NH_3^+$ and Na^+ sites are solvated becomes the most stable configuration for zwitterionic cysteine, and this completes the first solvent shell for Na^+Cys. A nearly identical structure around Na^+ is found for $Na^+Pro(H_2O)_4$, except the covalent N-C bond in Pro is replaced by a NH···S hydrogen bond in the Cys complexes. In both Cys and Pro complexes, the four water molecules interact electrostatically and directly with zwitterionic Na^+AA, whereas additional water molecules are likely to preferentially bind to these inner shell water molecules. In the glycine system, zwitterionic isomers in which both charge centers are solvated become favorable at about $x = 4$ (1 – 10 kJ/mol higher than the GS charge-solvated isomer), although until the first solvent shell is completed at five or six water molecules, the zwitterion won't be the clear ground state.[38] These comparisons indicate that intramolecular solvation provided by the functionalized side-chain allows more facile formation of the zwitterionic complex, which then requires a smaller first solvent shell (four waters) for stabilization of Na^+Cys versus that for Na^+Gly (five to six waters) because two additional waters are needed to completely solvate the NH_3^+ group of zwitterionic glycine. Because of this change, the hydration energies for Na^+Cys are found to exceed those for Na^+Gly and Na^+Pro, where there is less flexibility in the binding modes.

5. Conclusions

The kinetic energy dependences of the collision-induced dissociation (CID) of $Na^+Cys(H_2O)_x$, where $x = 1 - 4$ and Cys = cysteine, are examined in a guided ion beam tandem

mass spectrometer. The primary process observed in all cases is the loss of water from the complex. Sequential losses of water ligands are also observed for $x = 2 – 4$. The cross sections for losing cysteine from the complex are also observed at higher energies for $x = 1 – 4$ but are much less efficient than those for losing water. BDEs at 0 K for losing water from the complexes are measured from the threshold behavior and those for losing cysteine are derived from competitive modeling and thermodynamic cycles. The resulting experimental results show that the sequential binding energies for losing water or cysteine from $Na^+Cys(H_2O)_x$ decrease monotonically with increasing x. These trends are explained by increasing ligand-ligand repulsion and decreasing effective charge on the sodium ion as water molecules are added to the Na^+Cys complex.

Three different levels of quantum chemical calculations including zero point energy corrections and basis set superposition errors were performed for $Na^+Cys(H_2O)_x$, $x = 0 – 4$. Both experiment and calculations find the same general trends in the BDEs with increasing solvation, Fig. 7. The calculated BDEs for losing water and cysteine agree with our absolute experimental values within experimental uncertainties, although values for loss of Cys from $x = 3$ and 4 are towards the upper limits of experiment.

In the Na^+Cys complex, theory indicates that the sodium ion prefers to bind in a tridentate configuration to the carbonyl and amino group of the backbone and the thiol side-chain, ([N,CO,S] coordination).[30] The first and second water can attach directly to the sodium ion without influencing the cysteine binding conformation; however, at $x = 2$, DFT calculations suggest that it is isoenergetic for Na^+ to bind to the carbonyl with one of the water molecules bridging to the hydroxyl oxygen, Fig. 3. For $x = 3$, there are multiple isomers lying sufficiently low in energy that they could be populated in our experiment, which emphasizes that the third water can bind to Na^+, $–NH_3^+$, or –OH with nearly equal facility. For $x = 4$, the unique GS structure has a completely solvated tetracoordinate Na^+ and solvated $–NH_3^+$ group. Structures of this complex in which a water molecule binds to two inner shell water molecules to begin to form a second solvent shell are also low-lying.

UU IR Author Manuscript

When comparing $Na^+Cys(H_2O)_x$, x = 1 – 4, to the analogous glycine and proline systems,[38,39] we find that the BDEs for losing water are similar (within about 10 kJ/mol) but inversely track with the binding energy of Na^+ to the amino acid. The GS structures for $Na^+Cys(H_2O)_4$ and $Na^+Pro(H_2O)_4$ are very similar, which demonstrates that the strong hydrogen bond between the protonated amino group and the thiol side-chain provides strong intramolecular solvation of the $–NH_3^+$ charge site in zwitterionic Na^+Cys. Thus, the first hydration shells of Na^+Cys and Na^+Pro are complete with four water molecules and the amino acids in both species are zwitterionic. It seems likely that other amino acids with functionalized side chains should exhibit similar trends.

Acknowledgment. This work is supported by the National Science Foundation under Grant No. CHE-0748790. A grant of computer time from the Center for High Performance Computing in University of Utah is gratefully acknowledged.

References

[1] T. M. Larsen, L. T. Laughlin, H. M. Holden, I. Rayment, and G. H. Reed, *Biochemistry*, 1994, **33,** 6301.

[2] C. K. Mathews and K. E. van Holde, *Biochemistry* (Benjamin/Cummings Redwood City, 1990).

[3] C. H. Suelter, *Science*, 1970, **168,** 789.

[4] L. Stryer, *Biochemistry* (Freeman, New York, 1988).

[5] F. Jensen, *J. Am. Chem. Soc.*, 1992, **114,** 9533.

[6] G. Bojesen, T. Breindahl, and U. N. Andersen, *Org. Mass Spectrom.*, 1993, **28,** 1448.

[7] J. S. Klassen, S. G. Anderson, A. T. Blades, and P. Kebarle, *J. Phys. Chem.*, 1996, **100,** 14218.

[8] U. N. Andersen and G. Bojesen, *J. Chem. Soc. Perkins Trans. 2*, 1997**,** 323.

[9] B. A. Cerda, S. Hoyau, G. Ohanessian, and C. Wesdemiotis, *J. Am. Chem. Soc.*, 1998, **120,** 2437.

[10] S. Hoyau, K. Norrman, T. B. McMahon, and G. Ohanessian, *J. Am. Chem. Soc.*, 1999, **121,** 8864.

UU IR Author Manuscript

[11] T. Wyttenbach, M. Witt, and M. T. Bowers, *J. Am. Chem. Soc.*, 2000, **122,** 3458.

[12] B. A. Cerda and C. Wesdemiotis, *Analyst*, 2000, **125,** 657.

[13] V. Ryzhov, R. C. Dunbar, B. A. Cerda, and C. Wesdemiotis, *J. Am. Soc. Mass Spectrom.*, 2000, **11,** 1037.

[14] L. Rodriguez-Santiago, M. Sodupe, and J. Tortajada, *J. Phys. Chem. A*, 2001, **105,** 5340.

[15] T. Shoeib, A. C. Hopkinson, and K. W. M. Siu, *J. Phys. Chem. B*, 2001, **105,** 12399.

[16] T. Shoeib, K. W. M. Siu, and A. C. Hopkinson, *J. Phys. Chem. A*, 2002, **106,** 6121.

[17] R. M. Moision and P. B. Armentrout, *J. Phys. Chem. A*, 2002, **106,** 10350.

[18] T. Marino, N. Russo, and M. Toscano, *J. Phys. Chem. B*, 2003, **107,** 2588.

[19] M. M. Kish, G. Ohanessian, and C. Wesdemiotis, *Int. J. Mass Spectrom.*, 2003, **227,** 509.

[20] W. Y. Feng, S. Gronert, and C. B. Lebrilla, *J. Phys. Chem. A*, 2003, **107,** 405.

[21] C. Kapota, J. Lemaire, P. Maitre, and G. Ohanessian, *J. Am. Chem. Soc.*, 2004, **126,** 1836.

[22] R. M. Moision and P. B. Armentrout, *Phys. Chem. Chem. Phys.*, 2004, **6,** 2588.

[23] C. Ruan and M. T. Rodgers, *J. Am. Chem. Soc.*, 2004, **126,** 14600.

[24] R. M. Moision and P. B. Armentrout, *J. Phys. Chem. A*, 2006, **110,** 3933.

[25] S. J. Ye and P. B. Armentrout, *J. Phys. Chem. A*, 2008, **112,** 3587.

[26] S. J. Ye, A. A. Clark, and P. B. Armentrout, *J. Phys. Chem. B*, 2008, **112,** 10291.

[27] A. L. Heaton and P. B. Armentrout, *J. Phys. Chem. B*, 2008, **112,** 12056.

[28] A. L. Heaton, R. M. Moision, and P. B. Armentrout, *J. Phys. Chem. A*, 2008, **112,** 3319.

[29] P. B. Armentrout, A. Gabriel, and R. M. Moision, *Int. J. Mass Spectrom.*, 2009, **283,** 56.

[30] P. B. Armentrout, E. I. Armentrout, A. A. Clark, T. E. Cooper, E. M. S. Stennett, and D. R. Carl, *J. Phys. Chem. B*, 2010, **114,** 3927.

[31] V. N. Bowman, A. L. Heaton, and P. B. Armentrout, *J. Phys. Chem. B*, 2010, **114,** 4107.

[32] M. T. Rodgers and P. B. Armentrout, *Mass Spectrom. Rev.*, 2000, **19,** 215.

[33] M. T. Rodgers and P. B. Armentrout, *Accts. Chem. Res.*, 2004, **37,** 989.

[34] R. A. Jockusch, A. S. Lemoff, and E. R. Williams, *J. Phys. Chem. A*, 2001, **105,** 10929.

[35] R. A. Jockusch, A. S. Lemoff, and E. R. Williams, *J. Am. Chem. Soc.*, 2001, **123,** 12255.

[36] A. S. Lemoff, M. F. Bush, and E. R. Williams, *J. Am. Chem. Soc.*, 2003, **125,** 13576.

[37] A. S. Lemoff and E. R. Williams, *J. Am. Soc. Mass. Spectrom.*, 2004, **15,** 1014.

[38] S. J. Ye, R. M. Moision, and P. B. Armentrout, *Int. J. Mass Spectrom.*, 2005, **240,** 233.

[39] S. J. Ye, R. M. Moision, and P. B. Armentrout, *Int. J. Mass Spectrom.*, 2006, **253,** 288.

[40] A. S. Lemoff, M. F. Bush, and E. R. Williams, *J. Phys. Chem. A*, 2005, **109,** 1903.

[41] K. M. Ervin and P. B. Armentrout, *J. Chem. Phys.*, 1985, **83,** 166.

[42] K. M. Ervin and P. B. Armentrout, *J. Chem. Phys.*, 1984, **80,** 2978.

[43] F. Muntean and P. B. Armentrout, *J. Chem. Phys.*, 2001, **115,** 1213.

[44] E. R. Fisher and P. B. Armentrout, *J. Chem. Phys.*, 1991, **94,** 1150.

[45] E. R. Fisher, B. L. Kickel, and P. B. Armentrout, *J. Chem. Phys.*, 1992, **97,** 4859.

[46] R. H. Schultz, K. C. Crellin, and P. B. Armentrout, *J. Am. Chem. Soc.*, 1991, **113,** 8590.

[47] M. T. Rodgers and P. B. Armentrout, *J. Phys. Chem. A*, 1997, **101,** 1238.

[48] E. Teloy and D. Gerlich, *Chem. Phys.*, 1974, **4,** 417.

[49] D. Gerlich, *Adv. Chem. Phys.*, 1992, **82,** 1.

[50] N. R. Daly, *Rev. Sci. Instrum.*, 1960, **31,** 264.

[51] D. A. Hales, L. Lian, and P. B. Armentrout, *Int. J. Mass Spectrom. Ion Processes*, 1990, **102,** 269.

[52] M. T. Rodgers, K. M. Ervin, and P. B. Armentrout, *J. Chem. Phys.*, 1997, **106,** 4499.

[53] M. T. Rodgers and P. B. Armentrout, *J. Chem. Phys.*, 1998, **109,** 1787.

[54] R. G. Gilbert and S. C. Smith, *Theory of Unimolecular and Recombination Reactions* (Blackwell Scientific, London, 1990).

[55] T. S. Beyer and D. F. Swinehart, *Comm. Assoc. Computing Machinery*, 1973, **16,** 379.

[56] S. E. Stein and B. S. Rabinovich, *J. Chem. Phys.*, 1973, **58,** 2438.

[57] S. E. Stein and B. S. Rabinovich, *Chem. Phys. Lett.*, 1977, **49,** 183.

[58] E. V. Waage and B. S. Rabinovitch, *Chem. Rev.*, 1970, **70,** 377.

[59] P. B. Armentrout and J. Simons, *J. Am. Chem. Soc.*, 1992, **114,** 8627.

[60] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. R. Ross, T. E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, and P. Kollman, *Comp. Phys. Commun.*, 1995, **91,** 1.

[61] High Performance Computational Chemistry Group, Version 4.5 ed. (Pacific Northwest National Laboratory, Richland, Washington 99352, 2003).

UU IR Author Manuscript

[62] C. C. J. Roothan, *Rev. Mod. Phys.*, 1954, **23,** 69.

[63] J. S. Binkley, J. A. Pople, and W. J. Hehre, *J. Am. Chem. Soc.*, 1951, **102,** 939.

[64] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. Montgomery, J. A. , T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Revision B.02 ed. (Gaussian, Inc., Pittsburgh, PA, 2003).

[65] A. D. Becke, *J. Chem. Phys.*, 1993, **98,** 5648.

[66] R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.*, 1971, **72,** 5639.

[67] A. D. McLean and G. S. Chandler, *J. Chem. Phys.*, 1980, **72,** 5639.

[68] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, *J. Chem. Phys.*, 1980, **72,** 650.

[69] J. A. Montgomery, Jr., M. J. Frisch, J. W. Ochterski, and G. A. Petersson, *J. Chem. Phys.*, 1999, **110,** 2822.

[70] F. B. van Duijneveldt, J. G. C. M. van Duijneveldt de Rijdt, and J. H. van Lenthe, *Chem. Rev.*, 1994, **94,** 1873.

[71] P. B. Armentrout and M. T. Rodgers, *J. Phys. Chem. A*, 2000, **104,** 2238.

[72] C. H. S. Wong, F. M. Siu, N. L. Ma, and C. W. Tsang, *THEOCHEM*, 2002, **588,** 9.

[73] N. F. Dalleska, B. L. Tjelta, and P. B. Armentrout, *J. Phys. Chem.*, 1994, **98,** 4191.

[74] J. C. Amicangelo and P. B. Armentrout, *Int. J. Mass Spectrom.*, 2001, **212,** 301.

[75] T. E. Cooper, D. R. Carl, and P. B. Armentrout, *J. Phys. Chem. A*, 2009, **113,** 13727.

[76] S. J. Suresh and V. M. Naik, *J. Chem. Phys.*, 2000, **113,** 9727.

Table 1. Relative theoretical 0 K enthalpies (298 K free energies) in kJ/mol of $Na^+Cys(H_2O)_x$ [a]

complex	Structure	B3LYP	B3P86	MP2(full)
Na^+Cys	**[N,CO,S]-tggg$_+$**	**0.0 (0.0)**	**0.0 (0.0)**	**0.0 (0.0)**
	[N,CO,S]-tggg$_-$	4.6 (4.4)	3.4 (3.2)	4.1 (3.9)
	[CO_2^-]-ctgg$_+$	11.1 (9.3)	7.1 (5.2)	11.6 (9.8)
	[N,CO]-tcgg	15.3 (13.2)	14.4 (12.2)	16.4 (14.3)
	[N,OH,S]-tggg$_+$	30.6	31.9	27.2
	TS[CO_2^-/COOH]-ctgg$_+$	26.4	17.2	27.5
	[COOH]-ctgg$_+$	27.4	22.8	28.4
	[CO,S]-ctgt	23.7	21.5	31.6
$Na^+Cys(H_2O)$	**1W[N,CO,S]-tggg$_+$**	**0.0 (0.0)**	**0.0 (0.0)**	**0.0 (0.0)**
	1W[N,CO,S]-tggg$_-$	3.2 (3.6)	3.5 (3.9)	4.6 (5.0)
	1W[CO_2^-]-ctgg$_+$	8.9 (6.1)	4.9 (2.1)	11.9 (9.1)
	1W[N,CO]-tcgg	10.0 (7.5)	9.8 (7.3)	13.2 (10.8)
	1W[N,CO]-tgtg	13.0 (11.9)	12.6 (11.5)	17.0 (15.9)
	1W[bO^-]-[CO_2^-]-ctgg$_+$	15.6 (23.4)	9.2 (16.9)	17.9 (25.6)
	1W[HO]-[N,CO,S]-tggg$_+$	18.1 (23.9)	13.4 (19.2)	18.9 (24.7)
	1W[COOH]-ctgg$_+$	17.7 (13.6)	13.7 (9.5)	21.8 (17.6)
	TS-1W[CO_2^-/COOH]-ctgg$_+$	19.3 (16.7)	10.3 (7.7)	23.1 (20.5)
	1W[bCO]-[CO_2^-]-ctgg$_+$	21.4 (28.0)	15.4 (22.0)	23.3 (29.9)
	1W[bOH]-[CO]-ctgg$_+$	22.8 (27.9)	17.4 (22.6)	26.0 (31.1)
	1W[HN]-[CO_2^-]-ctgg$_+$	29.1 (30.5)	21.8 (23.2)	28.8 (30.1)
$Na^+Cys(H_2O)_2$	**2W[N,CO,S]-tggg$_+$**	**0.0 (0.0)**	5.9 (3.4)	**0.0 (0.0)**
	2W[N,CO,S]-tggg$_-$	6.0 (6.7)	9.9 (8.1)	5.3 (6.0)
	2W[HO]-[N,CO,S]-tggg$_+$	0.7 (2.7)	2.6 (2.1)	6.1 (8.0)
	2W[bO^-]-[CO_2^-]-ctgg$_+$	**0.0** (3.7)	**0.0** (1.2)	8.3 (12.0)
	2W[N,CO]-tcgg	3.5 (4.9)	9.8 (8.7)	9.8 (11.2)
	2W[bOH]-[CO]-ctgg$_+$	1.9 (2.4)	3.3 (1.4)	11.1 (11.7)
	2W[bOH]-[CO]-ctgg$_-$	1.9 (0.4)	4.0 **(0.0)**	11.3 (9.8)
	2W[bCO]-[CO_2^-]-ctgg$_+$	6.3 (9.0)	6.6 (6.8)	14.4 (17.1)
	TS-2W[bO^-/bOH]-[CO_2^-/CO]-ctgg$_+$	6.0 (9.9)	1.9 (3.4)	14.8 (18.7)
	2W[HN]-[CO_2^-]-ctgg$_+$	8.6 (4.1)	7.8 (0.8)	15.3 (10.8)
	2W[bOH,bS]-[CO]-ctgt	14.0 (20.9)	14.7 (19.1)	20.0 (26.8)

Table 1. continued

complex	Structure	B3LYP	B3P86	MP2(full)
$Na^+Cys(H_2O)_3$	**3W[bOH,bS]-[CO]-ctgt**	**0.0** (9.4)	3.0 (9.1)	**0.0** (6.2)
	3W[HO,bCO]-[N,CO,S]-tggg$_+$	4.6 (6.6)	7.7 (6.4)	1.2 **(0.0)**
	3W[bOH]-[CO]-ctgg$_+$	0.6 **(0.0)**	3.9 **(0.0)**	4.4 (0.6)
	3W[HN,bO$^-$]-[CO$_2^-$]-ctgg$_+$	1.6 (5.9)	**0.0** (1.0)	5.9 (7.0)
	3W[bO$^-$,bW]-[CO$_2^-$]-ctgg$_+$	4.9 (15.0)	5.4 (12.2)	7.2 (14.1)
	3W[bS,bCO,bW]-[N,CO]-tggg$_+$	12.2 (21.9)	18.4 (24.8)	7.4 (13.9)
	3W[bS,bCO,bW]-[N,CO]-tggt	15.8 (25.0)	21.6 (27.4)	8.0 (13.9)
	3W[HO,bS]-[N,CO]-tggg$_+$	7.0 (13.0)	9.7 (12.4)	8.6 (11.4)
	3W[bOH]-[CO]-cgtg	4.4 (6.1)	8.7 (7.1)	9.8 (8.2)
	3W[HN,bCO]-[CO$_2^-$]-ctgg$_+$	7.2 (10.1)	5.8 (5.4)	11.0 (10.6)
	TS-3W[bOH/bO$^-$,bW]-[CO/CO$_2^-$]-ctgg$_+$	9.1 (17.6)	5.7 (11.0)	11.6 (16.9)
$Na^+Cys(H_2O)_4$	**4W[HN,bO$^-$,bW]-[CO$_2^-$]-ctgg$_+$**	**0.0 (0.0)**	**0.0 (0.0)**	**0.0 (0.0)**
	4W[bOH,AA]-[CO]-ctgg$_+$	0.2 (2.5)	3.4 (5.7)	2.0 (4.3)
	4W[HN-bW,bO$^-$,AD]-[CO$_2^-$]-cgtg$_-$	9.7 (20.5)	7.9 (18.7)	4.1 (15.0)
	4W[HO,bS,bCO,bW]-[N,CO]-tggg$_+$	11.3 (15.7)	15.8 (20.2)	4.8 (9.2)
	4W[HN,bO$^-$,bW]-[CO$_2^-$]-cgtg$_-$	4.4 (4.3)	5.1 (4.9)	5.8 (5.7)
	4W[HN-bO$^-$r,bO$^-$,bWr]-[CO$_2^-$]-cgtg$_-$	9.2 (14.0)	10.2 (15.0)	6.1 (10.9)
	4W[HN-bO$^-$f,bO$^-$,bWf]-[CO$_2^-$]-cgtg$_-$	9.1 (13.3)	9.8 (13.9)	6.4 (10.6)
	4W[bOH,AA]-[CO]-cgtg$_-$	4.3 (6.2)	7.4 (9.3)	6.8 (8.6)
	4W[HN,bO$^-$,bW]-[CO$_2^-$]-ctgg$_-$	5.9 (5.6)	6.3 (5.9)	7.5 (7.1)
	4W[bO$^-$,AA]-[CO$_2^-$]-ctgg$_+$	6.6 (10.6)	7.7 (11.7)	7.5 (11.5)
	4W[bS,bCO,bHN,bW]-[N,CO]-tggg$_+$	23.0 (33.6)	27.8 (38.4)	7.8 (18.5)
	4W[HN-bW,bO$^-$,AD]-[CO$_2^-$]-cgtg$_+$	13.8 (23.5)	12.3 (22.0)	8.5 (18.2)
	4W[HN-bS,bO$^-$,bW]-[CO$_2^-$]-cgtg$_-$	10.7 (12.2)	10.7 (12.3)	8.9 (10.4)
	4W[bS,bOH,AD]-[CO]-ctgt	11.3 (19.8)	13.7 (22.1)	9.0 (17.4)
	4W[HO,bS,bW]-[N,CO]-tggt	18.4 (18.2)	21.4 (21.2)	9.1 (8.9)
	TS-4W[bOH/bO$^-$,AA]-[CO/CO$_2^-$]-ctgg$_+$	8.1 (13.2)	5.5 (10.5)	9.5 (14.6)
	4W[HN,bO$^-$,bW]-[CO$_2^-$]-cgtt	7.7 (6.3)	8.6 (7.3)	9.8 (8.4)
	4W[AA-bCO]-[CO$_2^-$]-ctgg$_+$	11.1 (18.0)	12.1 (18.9)	9.8 (16.7)

[a] Structures optimized at the B3LYP/6-311+G(d,p) level and all single point energies are calculated using the 6-311+G(2d,2p) basis set and the indicated level of theory. Ground state species at 0 and 298 K are indicated by bold face.

Table 2. Fitting parameters for eq 1 and entropies of activation at 1000 K[a]

Reactant	Ionic Product	d_J[b]	σ_0	n	E_0 (eV)	$\Delta S^{\ddagger}_{1000}$ (J/mol·K)
Na^+Cys[c]	Na^+	1	28 (8)	1.6 (0.3)	1.86 (0.14)	44 (2)
$Na^+Cys(H_2O)$	Na^+Cys	0.5	54 (5)	0.9 (0.1)	0.69 (0.09)	-12 (14)
	$Na^+(H_2O)$	0.5	2 (1)		1.42 (0.13)	23 (34)
$Na^+Cys(H_2O)_2$	$Na^+Cys(H_2O)$	0.5	97 (4)	0.9 (0.1)	0.56 (0.08)	18 (4)
	$Na^+(H_2O)_2$	0.5			1.05 (0.10)	64 (9)
$Na^+Cys(H_2O)_3$	$Na^+Cys(H_2O)_2$	0.5	117 (6)	1.0 (0.1)	0.46 (0.07)	27 (23)
	$Na^+(H_2O)_3$	0.33			0.82 (0.11)	90 (20)
$Na^+Cys(H_2O)_4$	$Na^+Cys(H_2O)_3$	0.5	63 (5)	1.2 (0.1)	0.35 (0.07)	35 (18)
	$Na^+(H_2O)_4$	0.25			0.68 (0.18)	109 (24)

[a] Uncertainties are listed in parentheses. [b] Reaction degeneracy: defined as the ratio of rotational symmetry numbers of the reactant to the products in the PSL TS. [c] Ref [30].

UU IR Author Manuscript

Table 3. Summary of energies (kJ/mol) for losing cysteine from $Na^+Cys(H_2O)_x$

Reactant	Competitive Fit[a]	Cycle 1[b]	Cycle 2[c]	Cycle 3[d]	Average[e]
Na^+Cys		177 (5)	**177 (5)**	169 (27)	177 (10)
$Na^+Cys(H_2O)$	137 (13)	148 (13)	148 (13)	140 (24)	144 (14)
$Na^+Cys(H_2O)_2$	101 (10)	109 (16)	120 (16)	112 (21)	107 (14)
$Na^+Cys(H_2O)_3$	79 (11)	75 (14)	94 (19)	86 (19)	81 (14)
$Na^+Cys(H_2O)_4$	65 (17)	58 (14)	73 (21)	**65 (17)**	64 (17)
MAD[f]	8 (5)	10 (8)	6 (5)	6 (4)	6 (5)

[a] Values from competitive modeling results in Table 2.

[b] Obtained from losing only single ligands, as described in the text, eq 5.

[c] Using $D_0(Na^+–Cys)$ (in bold) as the reference value, eq 6.

[d] Using $D_0(Na^+(H_2O)_4–Cys)$ (in bold) as the reference value, eq 7.

[e] Weighted average of all four values. Uncertainties are two standard deviations of the mean.

[f] Mean absolute deviation from the theoretical values in Table 5, excluding Na^+Cys.

Table 4. Enthalpies and free energies of H_2O and cysteine binding energies (kJ/mol) for $Na^+Cys(H_2O)_x$ (x = 0 – 4) at 0 and 298 K[a]

Complex	Ionic Product	ΔH_0[a]	$\Delta H_{298} - \Delta H_0$[b]	ΔH_{298}	$T\,\Delta S_{298}$[b]	ΔG_{298}
Na^+Cys	Na^+	177 (5)	1.6 (2.8)	179 (6)	33.8 (7.4)	145 (9)
$Na^+Cys(H_2O)$	Na^+Cys	66 (9)	-0.2 (0.5	66 (9)	26.9 (5.8)	39 (11)
	$Na^+(H_2O)$	144 (14)	-2.3 (0.4)	142 (14)	33.2 (6.0)	109 (15)
$Na^+Cys(H_2O)_2$	$Na^+Cys(H_2O)$	54 (8)	1.4 (2.2)	55 (8)	37.7 (7.8)	18 (11)
	$Na^+(H_2O)_2$	107 (14)	0.1 (2.3)	107 (14)	48.2 (9.3)	59 (17)
$Na^+Cys(H_2O)_3$	$Na^+Cys(H_2O)_2$	44 (7)	1.4 (4.2)	45 (8)	35.0 (14.7)	10 (17)
	$Na^+(H_2O)_3$	81 (14)	-0.8 (2.5)	80 (14)	47.6 (14.2)	33 (20)
$Na^+Cys(H_2O)_4$	$Na^+Cys(H_2O)_3$	34 (7)	3.1 (1.4)	37 (7)	41.4 (8.6)	-4 (11)
	$Na^+(H_2O)_4$	64 (17)	1.8 (1.0)	66 (17)	55.3 (6.9)	11 (18)

[a] Experimental values from this work (Tables 2 and 3). Uncertainties in parentheses.

[b] Values were computed using standard formulas and molecular constants calculated at the B3LYP/6-311+G(d,p) level. Uncertainties include 10% variations in the vibrational frequencies of the ligands and two-fold variations in the metal ligand-frequencies. Uncertainties also include variations in ground state conformers assumed for x = 2 and 3 as reactants and products.

UU IR Author Manuscript

Table 5. Experimental and theoretical H_2O and cysteine binding energies (kJ/mol) for $Na^+Cys(H_2O)_x$ at 0 K[a]

			Theory		
Complex	Ionic Product	Experiment	B3LYP	B3P86	MP2(full)
Na^+Cys	Na^+	177 (5) [b]	179.6	170.2	165.8
$Na^+Cys(H_2O)$	Na^+Cys	66 (9) [c]	65.5	63.2	64.7
	$Na^+(H_2O)$	144 (14) [d]	149.2	141.9	142.7
$Na^+Cys(H_2O)_2$	$Na^+Cys(H_2O)$	54 (8) [c]	45.1	49.1	48.4
	$Na^+(H_2O)_2$	107 (14) [d]	110.4	108.9	109.6
$Na^+Cys(H_2O)_3$	$Na^+Cys(H_2O)_2$	44 (7) [c]	45.2	43.8	42.2
	$Na^+(H_2O)_3$	81 (14) [d]	92.2	92.9	94.0
$Na^+Cys(H_2O)_4$	$Na^+Cys(H_2O)_3$	34 (7) [c]	38.7	40.5	35.3
	$Na^+(H_2O)_4$	64 (17) [d]	77.2	82.5	72.3
MAD[e]	$-H_2O$		4 (4)	4 (3)	3 (2)
	– Cys		7 (5)	8 (7)	7 (5)

[a] Energies calculated at the level indicated using the 6-311+G(2d,2p) basis set and B3LYP/6-311+G(d,p) geometries and vibrational frequencies. Zero point energies and BSSE corrections are included for all values. [b] Ref. [30]. [c] From competitive fitting in Table 2. [d] Weighted average values from Table 3. [e] Mean absolute deviation from the experimental values.

UU IR Author Manuscript

Figure Captions

Fig. 1. Zero pressure extrapolated cross sections for CID of Na^+CysW_x for $x = 1 – 4$ (parts a – d, respectively) where W = H_2O with Xe as a function of kinetic energy in the center-of-mass and laboratory frame. The solid lines show the model cross section convoluted over the neutral and ion kinetic and internal energies. The dashed lines show the model cross sections in the absence of experimental energy broadening for reactants with an internal energy of 0 K.

Fig. 2. Optimized structures of $Na^+Cys(H_2O)$ calculated at the B3LYP/6-311+G(d,p) level of theory. Relative 0 K energies in kJ/mol from single point energy calculations including zero point energies taken from Table 1 are indicated. Hydrogen bond lengths in Ångstroms are shown. (H – white; C – grey; N – blue; O – red; Na – purple; S – yellow)

Fig. 3. Optimized structures of $Na^+Cys(H_2O)_2$ calculated at the B3LYP/6-311+G(d,p) level of theory. Relative 0 K energies in kJ/mol from single point energy calculations including zero point energies taken from Table 1 are indicated. Hydrogen bond lengths in Ångstroms are shown. (H – white; C – grey; N – blue; O – red; Na – purple; S – yellow)

Fig. 4. Optimized structures of $Na^+Cys(H_2O)_3$ calculated at the B3LYP/6-311+G(d,p) level of theory. Relative 0 K energies in kJ/mol from single point energy calculations including zero point energies taken from Table 1 are indicated. Hydrogen bond lengths in Ångstroms are shown. (H – white; C – grey; N – blue; O – red; Na – purple; S – yellow)

Fig. 5. Optimized structures of $Na^+Cys(H_2O)_4$ calculated at the B3LYP/6-311+G(d,p) level of theory. Relative 0 K energies in kJ/mol from single point energy calculations including zero point energies taken from Table 1 are indicated. Hydrogen bond lengths in Ångstroms are shown. (H – white; C – grey; N – blue; O – red; Na – purple; S – yellow)

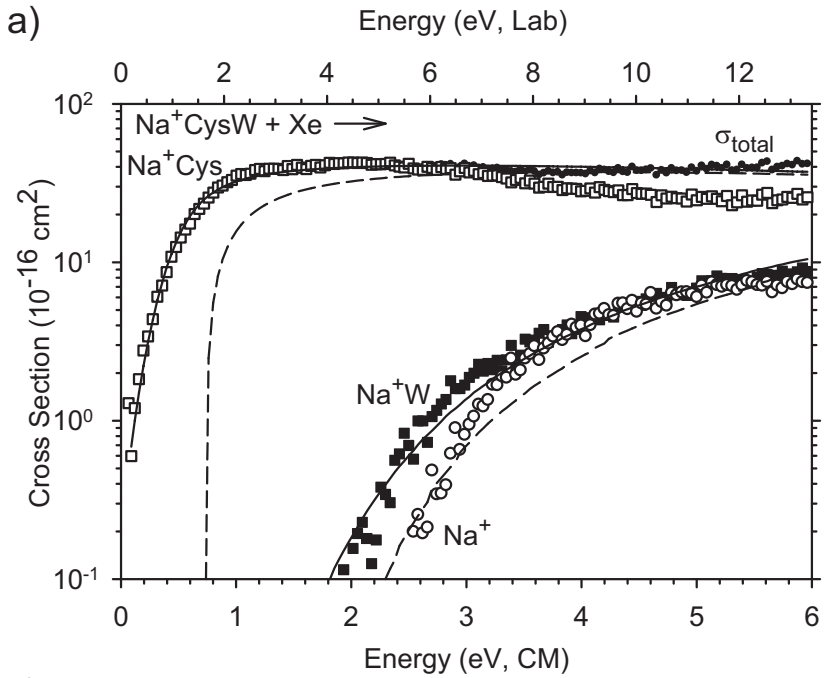
UU IR Author Manuscript

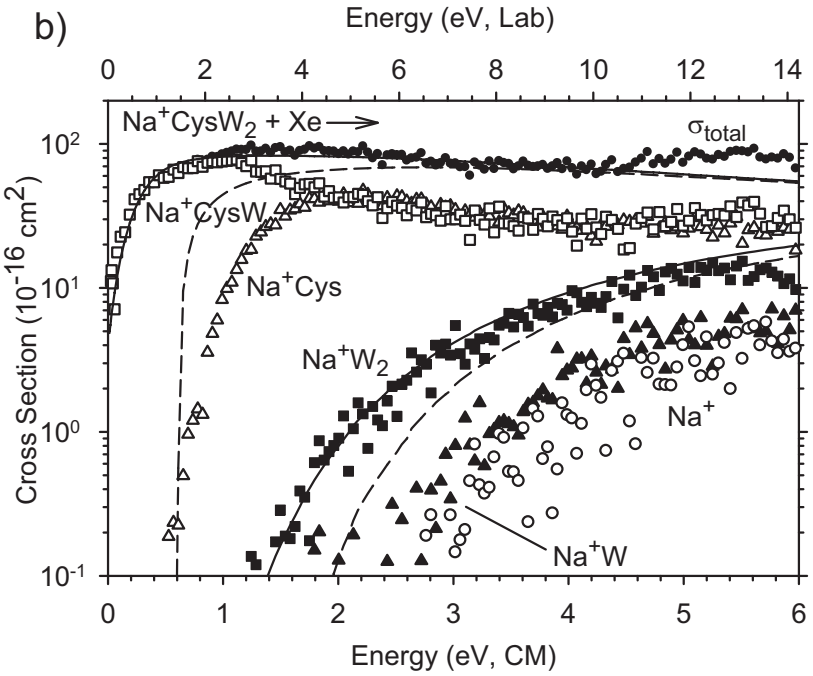
Fig. 6. Thermodynamic cycle for loss of Cys and H_2O from $Na^+Cys(H_2O)_x$, $x = 1 – 4$. All values in kJ/mol. Values for loss of Cys and H_2O from $Na^+Cys(H_2O)_x$ are taken from the competitive fits of experimental data in Table 2. Values for loss of H_2O from $Na^+(H_2O)_x$ are from Ref. [73].

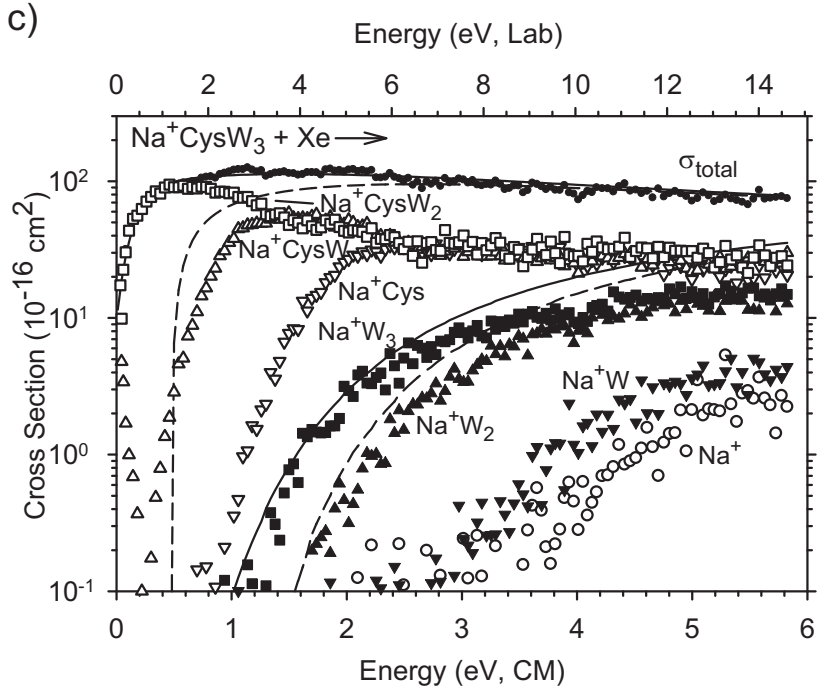
Fig. 7. Comparison of experimental (black open symbols) and theoretical (closed symbols: blue inverted triangles, B3LYP; green diamonds, B3P86; red triangles, MP2) 0 K BDE values in kJ/mol for loss of Cys (upper values) and water (lower values) from $Na^+Cys(H_2O)_x$ as a function of x, where $x = 0 – 4$. Values are from Table 5.

Fig. 8. Comparison of the 0 K BDEs (kJ/mol) for loss of the amino acid (AA) and water from $Na^+AA(H_2O)_x$ as a function of x, where $x = 0 – 4$ and AA = Cys (red circles, present data from Table 5), Pro (blue inverted triangles from Ref. [39]), and Gly (green triangles from Ref. [38]).

UU IR Author Manuscript


a)
Energy (eV, Lab)
Na+CysW + Xe
σtotal
Na+Cys
Na+W
Na+
Cross Section (10-16 cm2)
Energy (eV, CM)


b)
Energy (eV, Lab)
Na+CysW2 + Xe
σtotal
Na+CysW
Na+Cys
Na+W2
Na+
Na+W
Cross Section (10-16 cm2)
Energy (eV, CM)


c)
Energy (eV, Lab)
Na+CysW3 + Xe
σtotal
Na+CysW2
Na+CysW
Na+Cys
Na+W3
Na+W
Na+W2
Na+
Cross Section (10-16 cm2)
Energy (eV, CM)

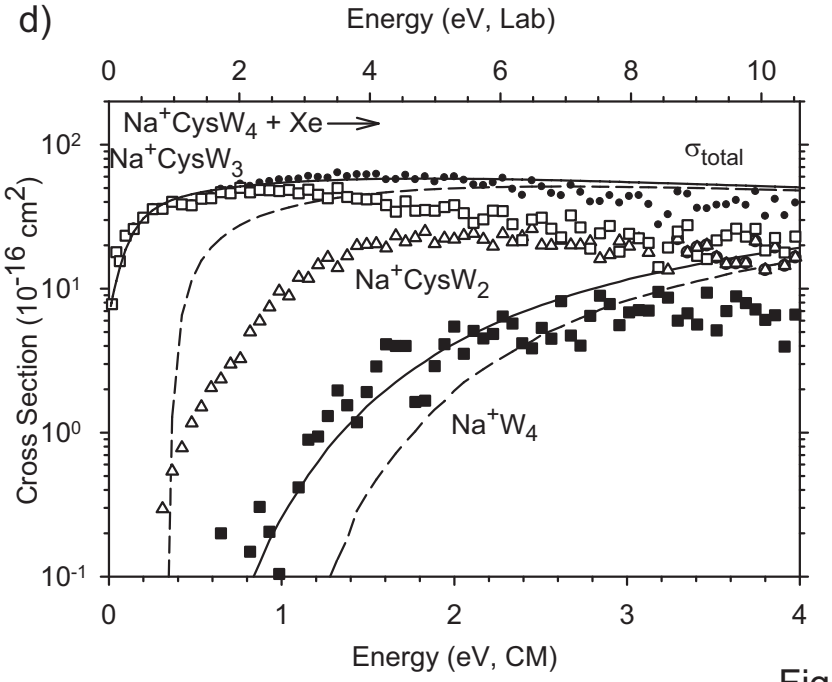
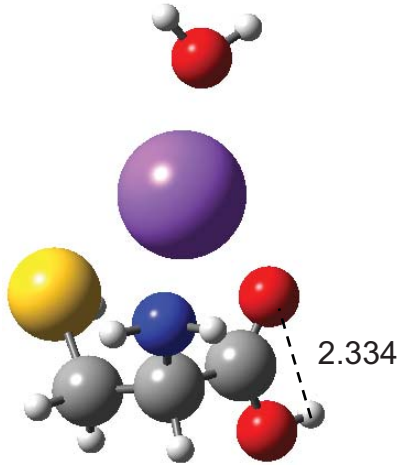
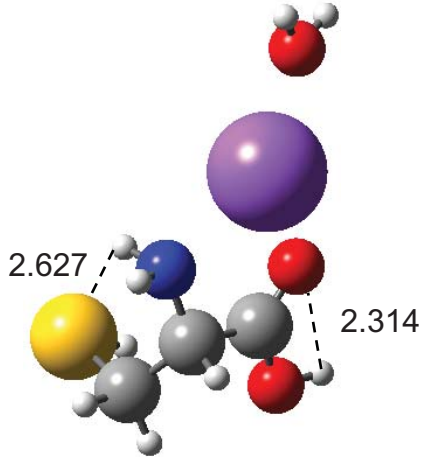

d)
Energy (eV, Lab)
Na+CysW4 + Xe
Na+CysW3
σtotal
Na+CysW2
Na+W4
Cross Section (10-16 cm2)
Energy (eV, CM)

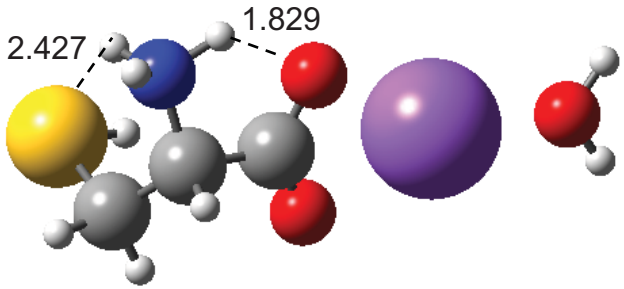
Figure 1



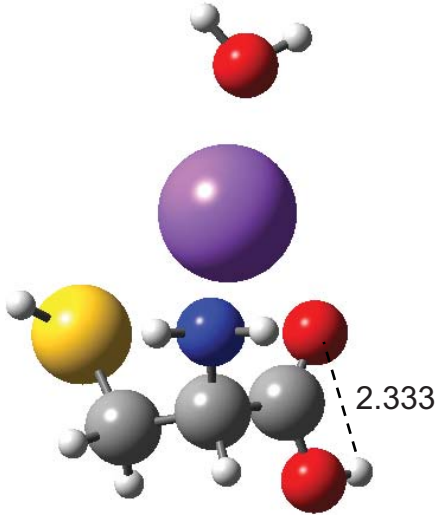
1W[N,CO,S]-tggg$_{+}$, 0



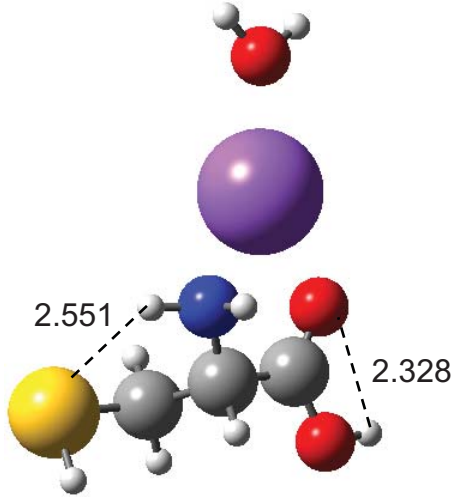
1W[N,CO]-tcgg, 10-13



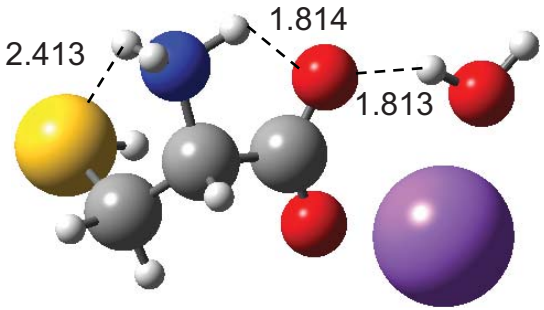
1W[CO_2^-]-ctgg$_{+}$, 5-12



1W[N,CO,S]-tggg$_{-}$, 3-5

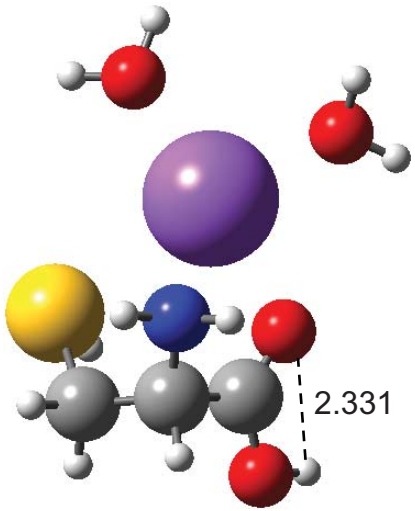


1W[N,CO]-tgtg, 12-17

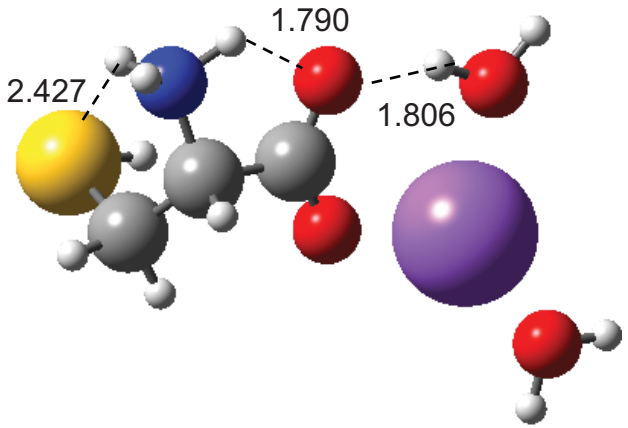


1W[bO$^-$]-[CO_2^-]-ctgg$_{+}$, 9-18

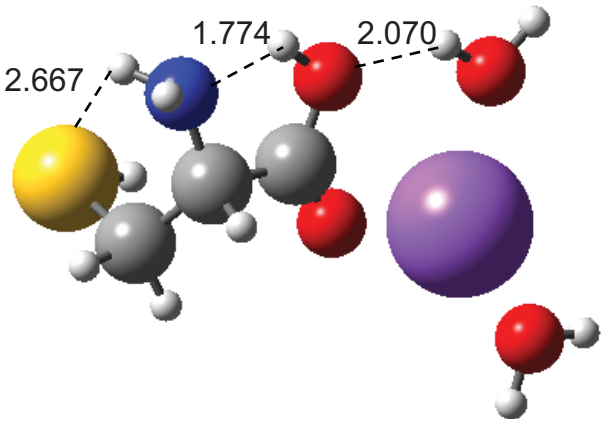
Figure 2



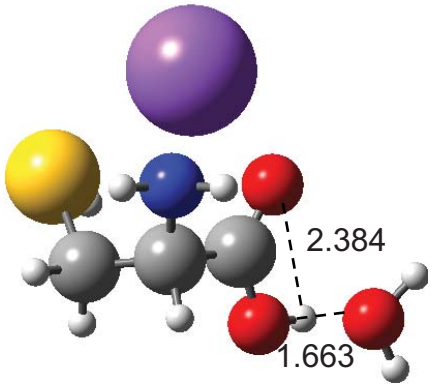
2W[N,CO,S]-tggg$_+$, 0-6



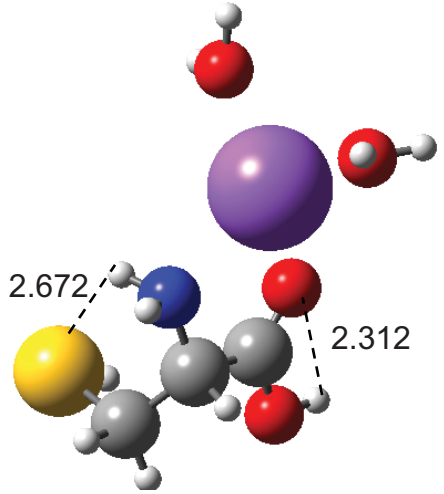
2W[bO$^-$]-[CO$_2^-$]-ctgg$_+$, 0-8



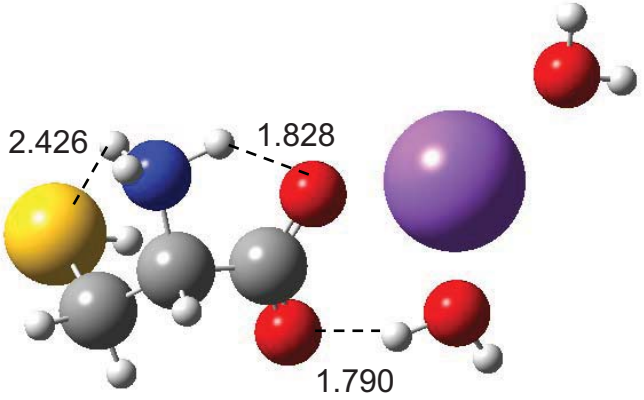
2W[bOH]-[CO]-ctgg$_+$, 2-11



2W[HO]-[N,CO,S]-tggg$_+$, 1-6

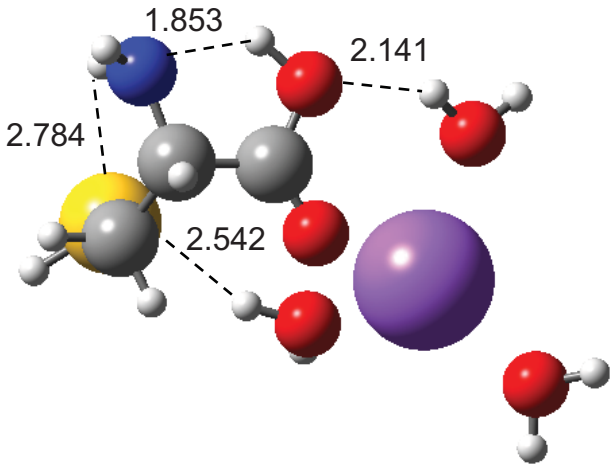


2W[N,CO]-tcgg, 3-10

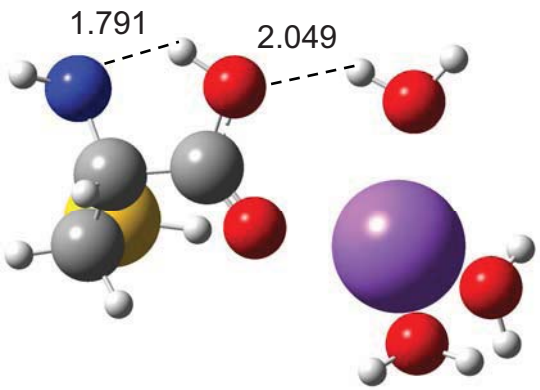


2W[bCO]-[CO$_2^-$]-ctgg$_+$, 6-14

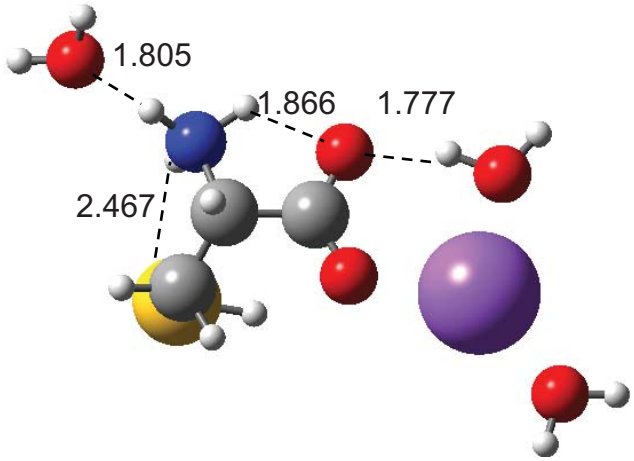
Figure 3



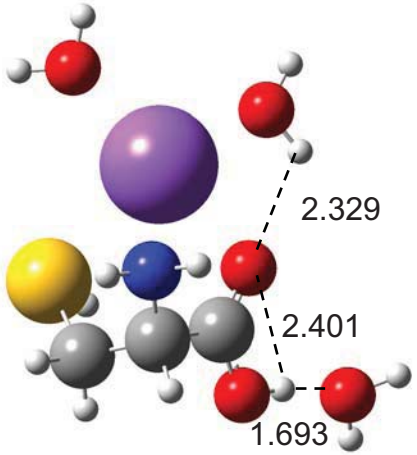
3W[bOH,bS]-[CO]-ctgt, 0-3



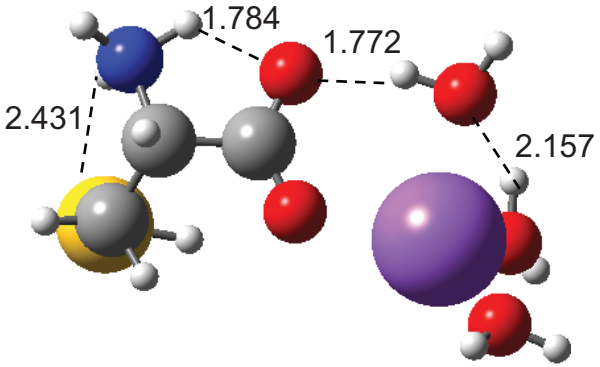
3W[bOH]-[CO]-ctgg$_+$, 0-4



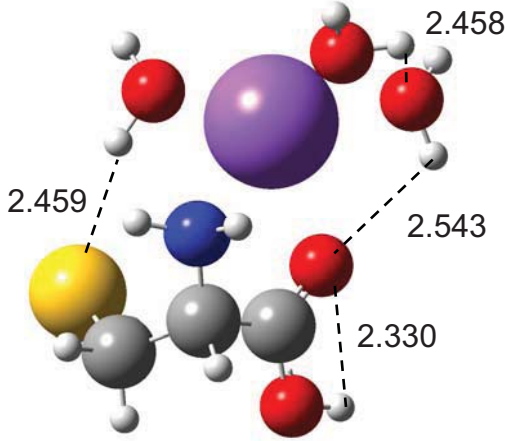
3W[HN,bO$^-$]-[CO$_2^-$]-ctgg$_+$, 0-6



3W[HO,bCO]-[N,CO,S]-tggg$_+$, 1-8

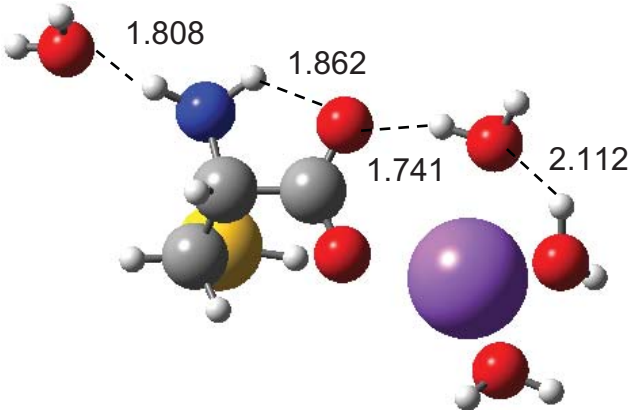


3W[bO$^-$,bW]-[CO$_2^-$]-ctgg$_+$, 5-7

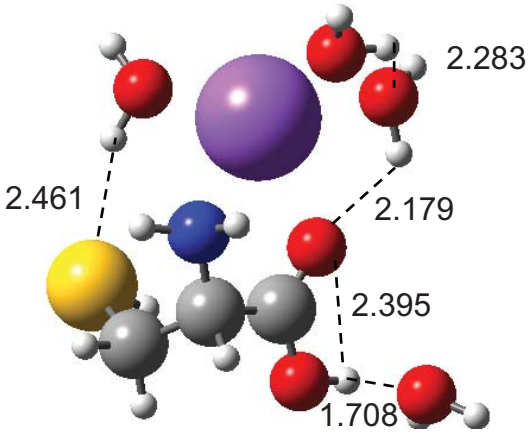


3W[bS,bCO,bW]-[N,CO]-tggg$_+$, 7-18

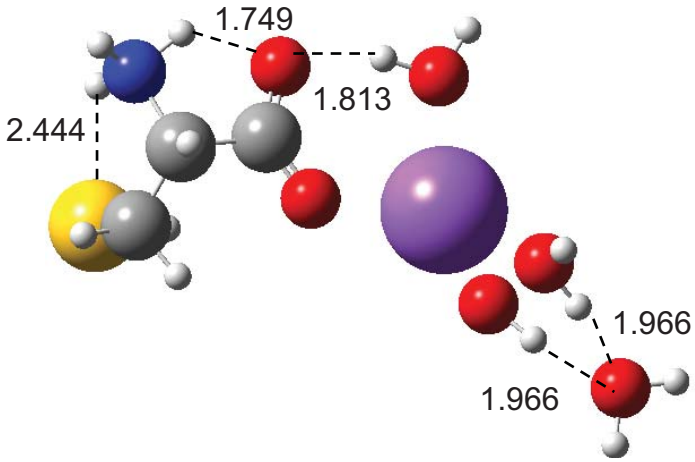
Figure 4



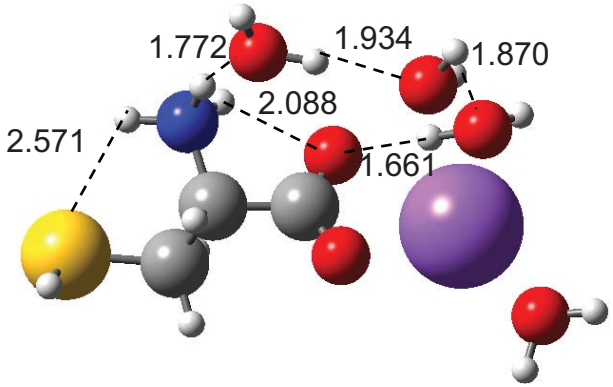
4W[HN,bO^{-},bW]-[CO_2^{-}]-ctgg$_{+}$, 0



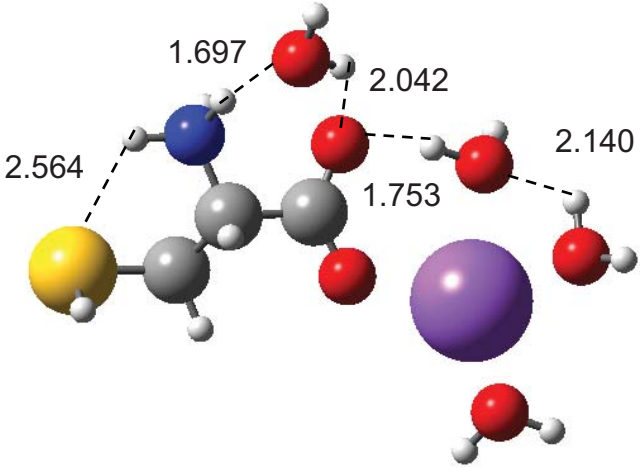
4W[HO,bS,bCO,bW]-[N,CO]-tggg$_{+}$, 5-16



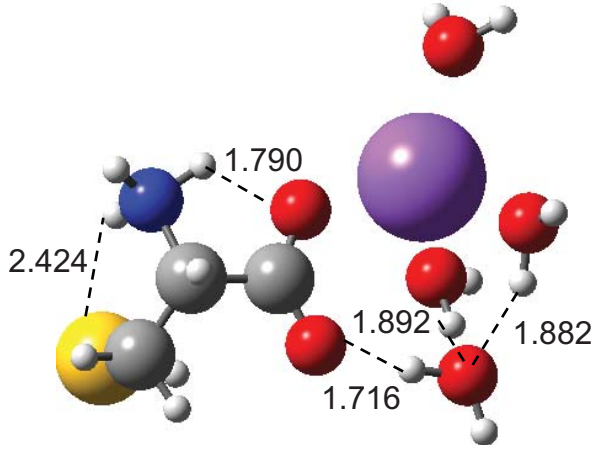
4W[bOH,AA]-[CO]-ctgg$_{+}$, 0-3



4W[HN-bW,bO^{-},AD]-[CO_2^{-}]-cgtg$_{-}$, 4-10



4W[HN-bO^{-}f,bO^{-},bWf]-[CO_2^{-}]-cgtg$_{-}$, 6-10



4W[AA-bCO]-[CO_2^{-}]-ctgg$_{+}$, 10-12

Figure 5

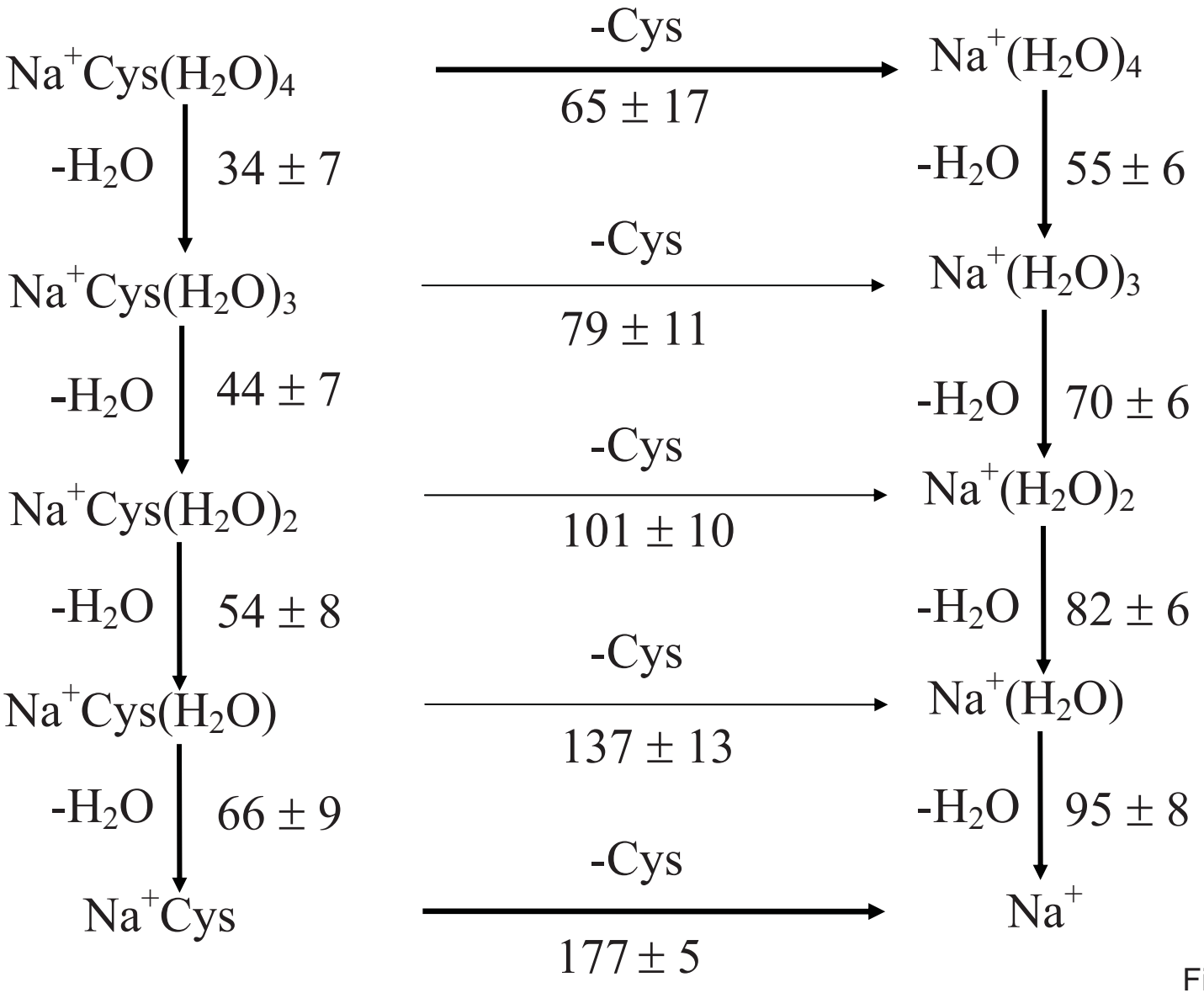

Na+Cys(H2O)4
-H2O
34 ± 7
Na+Cys(H2O)3
-H2O
44 ± 7
Na+Cys(H2O)2
-H2O
54 ± 8
Na+Cys(H2O)
-H2O
66 ± 9
Na+Cys
-Cys
65 ± 17
-Cys
79 ± 11
-Cys
101 ± 10
-Cys
137 ± 13
-Cys
177 ± 5
Na+(H2O)4
-H2O
55 ± 6
Na+(H2O)3
-H2O
70 ± 6
Na+(H2O)2
-H2O
82 ± 6
Na+(H2O)
-H2O
95 ± 8
Na+

Figure 6

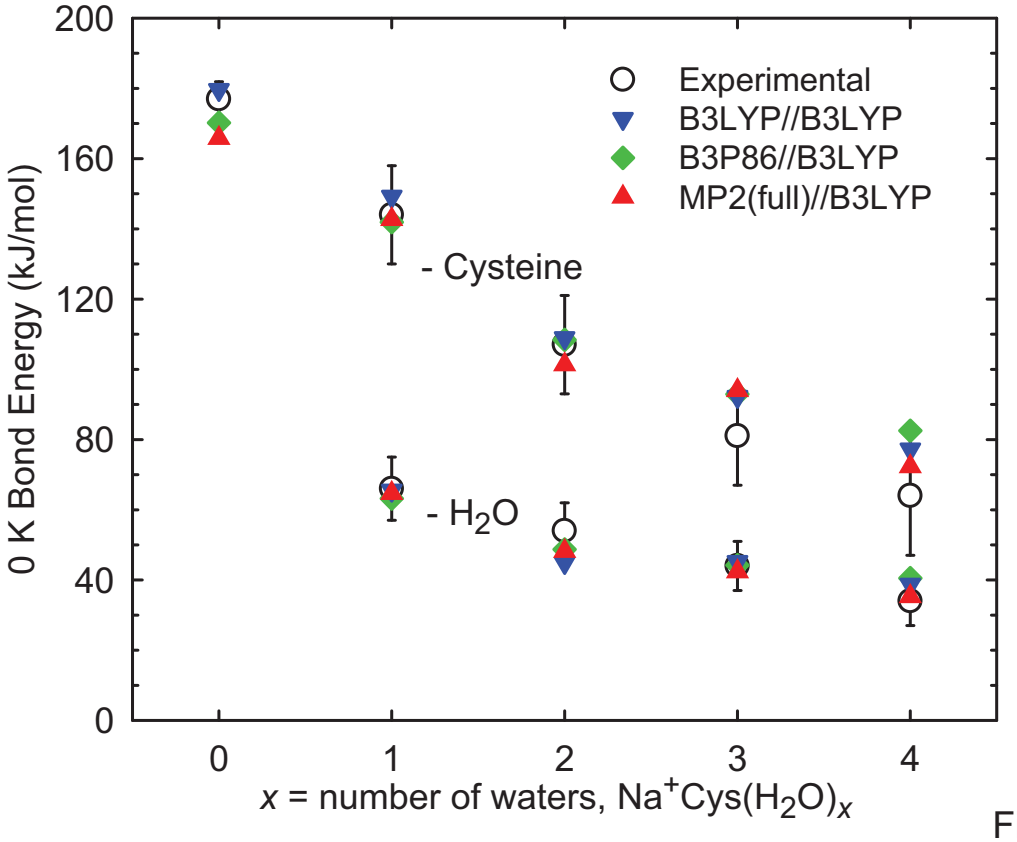


Figure 7

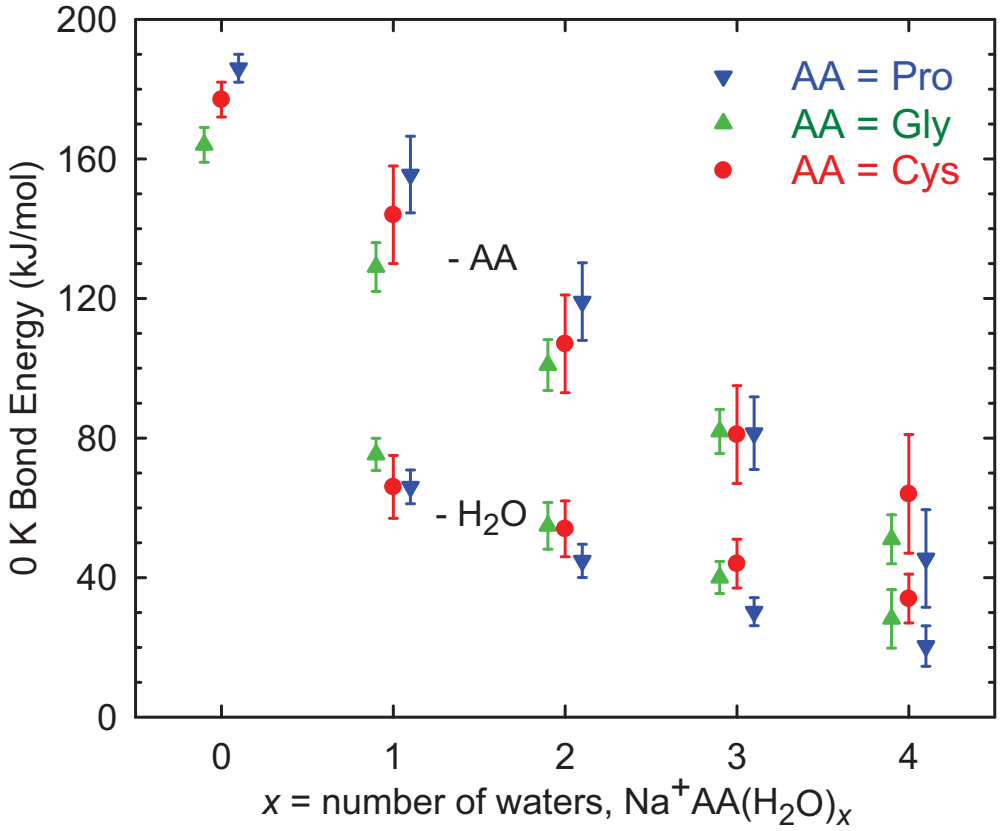


Figure 8

UU IR Author Manuscript